FIVE PROBLEMS
IN COMBINATORIAL NUMBER THEORY

BY

ZOLTÁN RÉTI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994

ACKNOWLEDGEMENTS

I would like to thank Dr. Frank Garvan for giving me the problem of Chapter 2, and for his guidance around t-cores.

I would like to thank three people for their time:

Dr. Frank Garvan (Chapter 2),

Dr. Gerhard Ritter (Chapter 5) and

Dr. Krishna Alladi (Chapter 1 and for being my advisor)

I would like to thank my father and mother for raising me with real love and support. My father filled the home of my childhood with books. He passed away two years ago in Budapest, without seeing me for the last time; I was in Gainesville. I would like to thank my brother too, who encouraged me to pursue mathematics.

TABLE OF CONTENTS

ABSTRACT v

CHAPTERS

1. INTRODUCTION 1

   1.1. The Problems 1
   1.2. Notations Used Everywhere 4

2. 3-CORE SOLUTIONS OF THE EQUATION $|x| + 2 \cdot |y| = n$ 5

   2.1. Introduction 5
   2.2. An Andrews-Garvan Crank for $s(3n+2)$ 8
   2.3. t-core Partitions 10
   2.4. Some Elementary Identities 12
   2.5. More Identities 15
   2.6. Various Expressions for the Number of t-cores 18
   2.7. A New Proof of the Cranks for $p_t^*(tn - (t^2-1)/24)$ 23
   2.8. An Identity of Nathan Fine's 30
   2.9. Two Cranks for $s^*(3n+2)$ 31
   2.10. Further Investigations; a Proof of $s^*(9n+8) \equiv 0 \pmod 9$ . . . 36

3. A CLASSIFICATION OF UNRESTRICTED PARTITIONS 42

4. AN EXTENSION OF BAILEY'S LEMMA 47

   4.1. The Classical Case 47
   4.2. The Extension 51
   4.3. Elementary Properties and Identities Used in the Proof 58

5. ORBITS OF ITERATED AFFINE TRANSFORMATIONS 61

   5.1. Introduction 61

5.2. The Solution Using the Jordan Normal Form 62
5.3. Projection of the Orbit, First Descent 65
5.4. Divided Differences. Some Notation 67
5.5. Projection of the Orbit, Full Descent 71
5.6. Finding the Orbit by Shortcutting the Descent 75
5.7. Computational Costs. Different Matrix Products 77
5.8. Dimension Two 80

6. DETERMINANTS, POWER SERIES, PARTITIONS 84

6.1. Introduction, Notation 84
6.2. Convolution of Sequences 86
6.3. Hessenberg Determinants and Combinatorial Numbers 92
6.4. Operations with Power Series Using Hessenberg Determinants . 96

REFERENCES 103

BIOGRAPHICAL SKETCH 106

Abstract of Dissertation Presented to
the Graduate School of the University of Florida
in Partial Fulfillment of the Requirements for the Degree
of Doctor of Philosophy

FIVE PROBLEMS
IN COMBINATORIAL NUMBER THEORY

By

Zoltán Réti

August 1994

Chairman: Dr. Krishna Alladi
Major Department: Mathematics

The common thread in the five chapters is that *partitions* of integers play at least marginal rôle in them. Three can be classified as belonging to number theory, more precisely partitions and basic hypergeometric series, one of them as belonging to combinatorics and one to linear algebra using some tools from approximation theory.

Problem One. Every partition of an integer n can be reduced to its unique t-core, where t is a fixed positive integer. The number of t-cores of n is related to an expression of η-functions. There are several ways (developed by F. Garvan, D. Stanton and others) to calculate this number. By introducing a new coordinate system, a simpler method is obtained.

A q-series identity of N. Fine's shows that the number of 3-core pair solutions (x, y) of $|x| + 2 \cdot |y| = 3n + 2$ is three times the number of solutions of $|x| + 2 \cdot |y| = n$. Surprisingly, two combinatorial statistics are found to explain this fact. The solutions are classified using both of them. This leads to a combinatorial explanation of a congruence of P. C. Eggans and an interesting self-similarity, among others.

There is also a new, constructive proof of the old fact that the number of t-cores of $tn + t - \frac{t^2-1}{24}$ can be divided by t, for $t = 5, 7, 11$.

Problem Two. As a complement to Dyson's rank, the notion of *frame* is introduced by the author to denote the (length of the largest part) + (number of parts) -1. Various properties of the numbers $p_r(n)$ (the number of partitions of n with frame size r) are noted, including a recursion formula. If the frame is even, then $p_r(n)$ is even, and this might help determine the parity of the number of partitions of n, $p(n)$, since $\sum_r p_r(n) = p(n)$. The sum $\sum_n p_r(n) = 2^{r-1}$ is also obtained.

Problem Three. The classical Bailey transform and the Bailey lemma contain a surprisingly large slice of the theory of the basic hypergeometric series. It was G. Andrews who recognized their significance and reformulated them as a matrix inversion problem. He iterated constant matrices to gain more and more complex identities and later, with A. Agarwal and D. Bressoud, was able to introduce some change in the matrices.

The generalization here can be regarded as a careful analysis and rethinking of their approach. The proof is admittedly very technical. The new parameters gained may lead to some new and useful identities.

Problem Four. In the course of solving the Problem One, affine transformations were used. Affine transformations are also important in several other fields such as fractal theory (image compression), robotics, etc. Recurrence sequences are further examples of affine transformations. Using some approximation theory, an algorithm and formulas are developed for iterated affine transformations, the results are decomposed for quicker computation in array computers. The special cases for the two dimensional transformations are also obtained.

Problem Five. The handling of convolutions of sequences is unified. This naturally leads to determinants of matrices whose entries above the superdiagonal are all

zeros. By analyzing their structure, it is shown how these determinants specialize to produce the major combinatorial numbers, such as the binomial coefficients, Stirling numbers of either kind.

A useful, partly survey section, with plenty of formulas, is also included about how one can manipulate formal power series using these determinants.

CHAPTER 1

INTRODUCTION

§1.1. The Problems

Dear Reader, you hold five chapters in your hand. Three can be classified as belonging to number theory, one of them as belonging to combinatorics and one to linear algebra using some tools from approximation theory.

So, what is common in them? Well, the common thread is *partitions* of integers. One chapter has this word in its title; another chapter deals with $t-core$ solutions of linear equations, and these $t-cores$ are special partitions, with t being a parameter. The topic of a third chapter is $q-series$, which is a device is to look at partition identities. Determinants look quite innocent, but the special ones we examine in chapter four can be expanded in the form $\sum c \cdot x_1^{j_1} \ldots x_d^{j_d}$, where $j_1 + \ldots + j_d =$ size of the determinant, i.e. it is the size that is being partitioned. What about the chapter on affine transformations? Here the main result (Theorem 1) contains summations over partitions of integers.

Let us emphasize that each chapter is self-contained with its own introduction; therefore, we will only sketch the main ideas of each chapter here.

3-core solutions of the equation $|x| + 2 \cdot |y| = n$. The concept of the *rank* of a partition [rank= largest part minus number of parts] was introduced by Dyson to give a combinatorial explanation to two of Ramanujan's congruences modulo 5 and 7 for the partition function. Dyson also pointed out that the rank does not explain Ramanujan's third (and deeper) congruence modulo 11, but he conjectured the existence of another statistic, which he called crank, to explain this third congruence.

Andrews and Garvan did find the crank [And88]. Since their work, there has been tremendous interest in statistics which would provide combinatorial explanations to other partition congruences. Such statistics are now given the generic name of cranks.

The following problem was raised by Frank Garvan. Every partition of an integer can be reduced to its unique t-core, where t is a fixed positive integer. Let $S^*(n)$ denote the set of pairs of 3-core solutions (x, y) of the equation $|x| + 2 \cdot |y| = n$, where $|x|$ is the partitioned integer. Using a q-series identity of N. Fine's, we see that $\#S^*(3n+2) = 3 \cdot \#S^*(n)$. Find a crank explaining this property combinatorially.

First, we introduce a new coordinate system, an injection A from the set of t-cores to $\mathbf{Z}^{t-1}$. It helps us find a very elegant proof of the old fact that the number of t-cores of $tn - \frac{t^2-1}{24}$ can be divided by t, for $t = 5, 7, 11$.

Next, surprisingly, we find two cranks for Garvan's problem. By using both cranks, we can classify the solutions in $S^*(n)$ into $t \times t$ matrices $M_t(n)$. These matrices have a lot of nice properties, including a self-similarity: the central ninth of $M_9(9n+8)$ is identical to $M_3(n)$.

We also prove that $M_3(9n+8)$ has equal entries, and this provides a combinatorial proof of a result of Eggan's [Egg89].

A Classification of the Partitions. As a complement to the rank, we introduce the term *frame (size)* to denote the (length of the largest part) + (number of parts) -1. In contrast to the rank, which has been studied in depth since Dyson first introduced it, comparatively little is known about the frame. Let $p_r(n)$ be the number of partitions of n with frame size r. Obviously $\sum_r p_r(n) = p(n)$, the number of partitions. We notice that $\sum_n p_r(n) = 2^{r-1}$ and show other properties including how to generate $p_r(n)$ from the set $\{p_j(n-r)\}_j$.

An extension of Bailey's lemma. The theory of partitions, started by Euler, advanced by many, most notably Sylvester and Ramanujan, leads to the investiga-

tions of q-series or basic hypergeometric series in a natural way. The classical Bailey transform and the Bailey lemma contain a surprisingly large slice of the theory of these series. It was George Andrews who recognized their significance later and reformulated them as a matrix inversion problem. He iterated constant matrices to gain more and more complex identities. Bressoud et al. [Aga87] further modified the method by changing the matrices at each iteration.

Our generalization can be regarded as a careful analysis and rethinking of their approach. The proof is admittedly very technical. Using the formula, we can reduce a multiple sum into a double sum. In the classical case the double sum is further reduced to a single sum, which in turn can yield a product. We gain new parameters. On one hand, our analysis helps us better understand the mechanism of the Bailey chains; on the other hand, the new parameters may lead to some new and useful identities.

Orbits of iterated affine transformations. In the course of solving the problem on t-cores, we used affine transformations. Affine transformations are also important in several other fields such as fractal theory (image compression), robotics, etc. Recurrence sequences are further examples of affine transformations. Using some *approximation theory,* we develop an algorithm and formulas for iterated affine transformations, then we decompose the results for quicker computation in array computers.

Determinants, power series, partitions. In this chapter we unify the handling of convolutions of sequences. This naturally leads to determinants of matrices whose entries above the superdiagonal are all zeros. By analyzing their structure, we show how these determinants specialize to produce the major combinatorial numbers, such as the binomial coefficients, Stirling numbers of either kind. We also include a partly survey section about how easily we can manipulate formal power series using them.

§1.2. Notations Used Everywhere

We will use // to extract a coefficient from a formal power series,i.e. $\sum c_k q^k // q^n := c_n$. This notation has certain advantages, e.g. we can simplify: $\sum c_k q^k // q^n = \sum c_k q^{k-m} // q^{n-m}$.

The next notation is standard in the literature.

$$(x)_n := (x;q)_n := \begin{cases} (1-x)(1-xq)\cdot \ldots \cdot (1-xq^{n-1}) & \text{if } n > 0 \\ 1 & \text{if } n = 0 \\ 1/\left((1-xq^{-1})(1-xq^{-2})\cdot \ldots \cdot (1-xq^{n})\right), & \text{if } n < 0. \end{cases}$$

Following [Gas90], we shall use the abbreviation

$(x_1, x_2, \ldots, x_k; q^t)_n := (x_1, q^t)_n (x_2, q^t)_n \cdot \ldots \cdot (x_k, q^t)_n$

The notations $a := b$ and $b =: a$ both mean that a is defined by b.

Finally, sets and matrices are capitalized, while vectors are underlined, e.g. $\underline{v}$.

CHAPTER 2

3-CORE SOLUTIONS OF THE EQUATION $|x| + 2 \cdot |y| = n$

§2.1. Introduction

A *partition* π of n is a non-increasing sequence of positive integers whose sum is n. We will write $|\pi| = n$ to express this relationship. For example, $|5+5+4+2+1+1| = 18$. Let $P(n)$ denote the set and let $p(n)$ denote the number of partitions of n. Ramanujan proved that

$$p(5n+4) \equiv 0 \pmod 5 , \tag{1}$$

$$p(7n+5) \equiv 0 \pmod 7 , \tag{2}$$

and

$$p(11n+6) \equiv 0 \pmod{11} \tag{3}$$

hold.

Dyson [Dys44] introduced the notion of the *rank* to interpret these congruences combinatorially. The rank is just the largest part minus the number of parts. For instance, $rank(5+5+4+2+1+1) = 5 - 6 = -1$.

Let $p\,(n, rank \equiv k \text{ (mod t)})$ denote the number of partitions of n whose rank is congruent to k modulo t. Dyson conjectured, Atkin and Swinnerton-Dyer [Atk54] proved that

$$p\,(5n+4, rank \equiv k \text{ (mod 5)}) = \frac{p(5n+4)}{5} \text{ for each } k,\ 0 \le k \le 4, \tag{4}$$

and

$$p\,(7n+5, rank \equiv k \text{ (mod 7) }) = \frac{p(7n+5)}{7} \text{ for each } k,\ 0 \le k \le 6. \tag{5}$$

The analogue of statements (4) and (5) for $t = 11$ is false. It was more than thirty years later, when Andrews and Garvan [And88] found a new statistic, called the *crank*, which gives

$$p\,(11n+6, crank \equiv k \text{ (mod 11) }) = \frac{p(11n+6)}{11} \text{ for each } k,\ 0 \le k \le 10. \tag{6}$$

Formulas (4) and (5) are also satisfied if *rank* is replaced by *crank*. The crank is just the largest part, if there are no 1's in the partition, and the number of parts that are larger than μ minus μ, if μ, the number of 1's in the partition is greater than zero. For instance, $crank(5+5+4+2+1+1) = 3-2 = 1$, $crank(5+5+4+2+1) = 4-1 = 3$ and $crank(5+5+4+2) = 5$. There is no known combinatorial proof of (4),(5) or (6) using the rank or this crank.

But Garvan, Kim and Stanton [Gar90] did find a combinatorial proof using a *different* crank. They regarded *t-core* partitions of n. (We will define and examine them in detail in the forthcoming sections). Let $p_t^*(n)$ denote the number of $t-core$ partitions of n. Congruences (1), (2) and (3) remain valid when $p(n)$ is replaced by $p_t^*(n)$. They found three new cranks, one for each $t = 5, 7, 11$, such that

$$p_t^*\left(tn - \frac{t^2-1}{24}, crank_t \equiv k \text{ (mod t)}\right) = \frac{p_t^*(tn - \frac{t^2-1}{24})}{t} \text{ for each } k,\ 0 \le k \le t-1. \tag{7}$$

These cranks are defined on the set of *t-cores,* denoted by $P_t^*(n)$. But due to the identities $p_t^*(n) = \frac{(q^t;q^t)^t_\infty}{(q)_\infty}//q^n$ and $p(n) = \frac{1}{(q)_\infty}//q^n$, these cranks extend to the set of all partitions $P(n)$, proving (4), (5) and (6). We will give a new, constructive proof of (7) at the end of the chapter, but our thrust is another problem.

Let us note here that (1), (2) and (3) have the following generalizations, due to Watson [Wat38] and Atkin [Atk67]:

$$p(5^\alpha n + \delta_{5,\alpha}) \equiv 0 \pmod{5^\alpha}\ , \tag{1'}$$

$$p(7^\alpha n + \delta_{7,\alpha}) \equiv 0 \pmod{7^{1+\lfloor \frac{\alpha}{2} \rfloor}}\ , \tag{2'}$$

and

$$p(11^\alpha n + \delta_{11,\alpha}) \equiv 0 \pmod{11^\alpha}\ , \tag{3'}$$

where

$$24\delta_{t,\alpha} \equiv 1 \pmod{t^\alpha},\ t = 5, 7, 11. \tag{8}$$

The analogous generalizations for *t-core* partitions were proven by Garvan [Gar93]:

$$p_t^*\left(t^\alpha n - \frac{t^2-1}{24}\right) \equiv 0 \pmod{t^\alpha}\ , \tag{9}$$

where $t = 5, 7, 11$.

Let us introduce the main problem of this section. Eggan [Egg89] studied the numbers

$$s(n) := \frac{1}{(q)_\infty} \cdot \frac{1}{(q^2;q^2)_\infty} // q^n. \tag{10}$$

He found that they satisfy the congruences

$$s(n) \equiv 0 \pmod{3^\alpha},\ \text{whenever } 8n \equiv 1 \pmod{3^\alpha}\ . \tag{11}$$

We will give two kinds of cranks which interpret (11) combinatorially for $\alpha = 1, 2$. The first crank, found by Garvan, is similar to the Andrews-Garvan crank and proves (11) with $\alpha = 1$.

The second kind has the Garvan-Kim-Stanton flavor of extension. We will examine the coefficients

$$s^*(n) := \frac{(q^3;q^3)^3_\infty}{(q)_\infty} \cdot \frac{(q^6;q^6)^3_\infty}{(q^2;q^2)_\infty} // q^n, \tag{12}$$

and, surprisingly, we will able to find two cranks; their extension combinatorially proves (11) with $\alpha = 1, 2$ (see§8 − 10). The title of the chapter comes from the fact that $s^*(n)$ is the number of 3-core pairs (x, y) satisfying $|x| + 2|y| = n$.

We will start with the first kind of crank.

§2.2. An Andrews-Garvan Crank for $s(3n + 2)$

The results of this section are due to Garvan (personal communication). Let ω be a t^{th} root of unity. The Andrews-Garvan crank depended on the analysis of the product $\frac{(q)_\infty}{(\omega)_\infty(\omega^{-1})_\infty}$. We find a crank by introducing the cube root of unity ω in the product

$$\frac{1}{(q)_\infty} \cdot \frac{1}{(q^2;q^2)_\infty}.$$

Since this product is equal to

$$\prod_{n=1}^{\infty} \frac{1}{(1-q^{2n})^2 \cdot (1-q^{2n-1})} = \frac{1}{(q^2, q^2, q; q^2)_\infty}, \tag{13}$$

we can interpret $s(n)$ as the number of partitions of n in which the even parts may have two colors (denoted by subscripts 0 and 1). For example, $s(5) = 12$ and the partitions are $\{5, 4_1 + 1, 4_0 + 1, 3 + 2_1, 3 + 2_0, 3 + 1 + 1, 2_1 + 2_1 + 1, 2_1 + 2_0 + 1, 2_0 + 2_0 + 1, 2_1 + 1 + 1 + 1, 2_0 + 1 + 1 + 1, 1 + 1 + 1 + 1 + 1\}$.

We define the crank for the $s(3n+2) \equiv 0 \pmod 3$ problem (which is (11) with $\alpha = 1$) to be the number of even parts colored 1 minus the number of even parts colored 0.

Let $s\,(n, crank \equiv k \text{ (mod 3) })$ denote the number of colored partitions with the crank defined above being congruent to k (mod 3) and let $s(n, crank = m)$ denote the number of colored partitions enumerated by $s(n)$ with crank equalling m. Obviously, for $k = 0, 1, 2$ we have

$$s\,(n, crank \equiv k \text{ (mod 3) }) = \sum_{m \equiv k (mod 3)} s(n, crank = m). \tag{14}$$

THEOREM 1. *For k=0,1,2 we have*

$$s\,(3n+2, crank \equiv k \ (mod\ 3)\ ) = \frac{s(3n+2)}{3}.$$

PROOF: Observe that

$$s(n, crank = m) = \frac{1}{(q, zq^2, z^{-1}q^2; q^2)_\infty} // q^n z^m \tag{15}$$

Substitute $z = \omega$, and use (14) to get

$$s\,(n, crank \equiv k \text{ (mod 3) }) = \frac{1}{(q, \omega q^2, \omega^2 q^2; q^2)_\infty} // q^n \omega^k. \tag{16}$$

Take $x := q^2$, then substitute q^2 in the place of q in the identity

$$(x, x\omega, x\omega^2)_\infty = (x^3; q^3)_\infty \tag{17}$$

to rewrite the generating function of (16) as

$$\frac{1}{(q;q^2)_\infty} \cdot \frac{1}{(\omega q^2, \omega^2 q^2; q^2)_\infty} = \frac{1}{(q;q^2)_\infty} \cdot \frac{(q^2;q^2)_\infty}{(q^6;q^6)_\infty}.$$

But

$$\frac{(q^2;q^2)_\infty}{(q;q^2)_\infty} = \sum_{j=0}^{\infty} q^{\frac{j(j+1)}{2}} \tag{18}$$

by a lemma of Gauss' (see [And76, page 23]), so

$$s\,(n, crank \equiv k \pmod 3\,) = \frac{\sum_{j=0}^{\infty} q^{\frac{j(j+1)}{2}}}{(q^6;q^6)_\infty} // q^n\omega^k, \text{ by (16)}\ .$$

Therefore

$$\sum_{k=0}^{2} s\,(n, crank \equiv k \pmod 3\,)\,\omega^k = \frac{\sum_{j=0}^{\infty} q^{\frac{j(j+1)}{2}}}{(q^6;q^6)_\infty} // q^n. \tag{19}$$

Suppose $n \equiv -1 \pmod 3$. In this case $\frac{j(j+1)}{2} + 6m \not\equiv -1 \equiv n$ shows that the right hand side of (19) is 0. From this we can conclude that

$$\begin{aligned} s\,(3n+2, crank \equiv 0 \pmod 3\,) &= s\,(3n+2, crank \equiv 1 \pmod 3\,) \\ &= s\,(3n+2, crank \equiv 2 \pmod 3\,)\,, \end{aligned} \tag{20}$$

which gives the statement of the theorem. ∎

We will find two other cranks for $s(3n+2)$ in §9, but we have to build our tools first.

§2.3. t-core Partitions

Write the Ferrers diagram of a partition π on a transparency and place it over a $t-$colored chessboard. Make sure the upper left corners of the tranparency and the chessboard coincide and color the dots of the Ferrers diagram by the color of the underlying chessboard. The colors will be denoted by $0, 1, \ldots, t-1$, occurring in this order in the first row, from left to right. In general, the dot at (row r, column c)

will assume color $c - r \pmod t$. The resulting colored Ferrers graph is called the *t-residue diagram* of the partition. For example, the Ferrers diagram of partition $6+6+5+5+5+3+1+1$ is

$$
\begin{array}{cccccc}
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \\
\bullet & \bullet & \bullet & \bullet & \bullet & \\
\bullet & \bullet & \bullet & \bullet & \bullet & \\
\bullet & \bullet & \bullet & & & \\
\bullet & & & & & \\
\bullet & & & & &
\end{array}
$$

and its coloring by the 2-, 3- and 4-color chessboards results in diagrams

$$
\begin{array}{cccccc}
0 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & \\
1 & 0 & 1 & 0 & 1 & \\
0 & 1 & 0 & 1 & 0 & \\
1 & 0 & 1 & & & \\
0 & & & & & \\
1 & & & & &
\end{array}
\qquad
\begin{array}{cccccc}
0 & 1 & 2 & 0 & 1 & 2 \\
2 & 0 & 1 & 2 & 0 & 1 \\
1 & 2 & 0 & 1 & 2 & \\
0 & 1 & 2 & 0 & 1 & \\
2 & 0 & 1 & 2 & 0 & \\
1 & 2 & 0 & & & \\
0 & & & & & \\
2 & & & & &
\end{array}
\qquad
\begin{array}{cccccc}
0 & 1 & 2 & 3 & 0 & 1 \\
3 & 0 & 1 & 2 & 3 & 0 \\
2 & 3 & 0 & 1 & 2 & \\
1 & 2 & 3 & 0 & 1 & \\
0 & 1 & 2 & 3 & 0 & \\
3 & 0 & 1 & & & \\
2 & & & & & \\
1 & & & & &
\end{array}
$$

respectively. For a fixed t, let $r_j = r_j(\pi)$ be the number of dots colored j. In the example above, for $t = 4$ we have $r_0 = 9$, $r_1 = 9$, $r_2 = 7$ and $r_3 = 7$. From each dot of the diagram we can make eastward and southward cuts. The Γ-shaped cuts obtained are called *hooks.* The crooked segments of the border which connect the eastern and southern tips of the Γ's are the *rimhooks.* Using the previous example, let us initiate the cut at dot $(2,2)$ and mark the dots belonging to the hook, rimhook or both by h, r and b. We obtain

```
. . . . . .
. h h h b b
. h . . r
. h . . r
. h r r r
. b r
.
.
```

Let $h(x,y)$ denote the number of dots belonging to the hook originated at dot (x,y) and call it *hooklength* ; in our example $h(2,2) = 9$. A hook and the corresponding rimhook have the same number of dots.

The $t-$ *weight* of π is the number of dots (x,y) such that t divides $h(x,y)$. If there is no such a point, i.e. the t-weight is zero, the partition is called a *t-core*.

If we are given a partition and find a hook with $h = k \cdot t$, we can remove the corresponding rimhook from the diagram. In the remaining diagram, the number of dots colored j, r'_j satisfies $r'_j = r_j - k$. The repeated removing process will stop; and no matter in which order we removed the rimhooks whose length was divisible by t, we end up with the same t-core. This t-core is, of course, the partition of a smaller integer. If π_1 and π_2 are partitions of the same integer, then their t-core is the same iff $r_j(\pi_1) = r_j(\pi_2)$ for every $j = 0, 1, \ldots, t-1$ [Jam81, page 87].

In the next two sections we leave, seemingly, the realm of t-cores, but we return in force in §6, armed with the results of the investigations of §4 and §5.

§2.4. Some Elementary Identities

Although we defined numbers r_j for a partition π as the number of dots colored

j by the underlying t-colored chessboard, the results of this section are valid for arbitrary complex numbers r_j. Let

$$\underline{r} := \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{t-1} \end{bmatrix},$$

and define

$$\underline{c} := \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{t-2} \\ c_{t-1} \end{bmatrix} := \begin{bmatrix} r_0 - r_1 \\ r_1 - r_2 \\ \vdots \\ r_{t-2} - r_{t-1} \\ r_{t-1} - r_0 \end{bmatrix} \text{ and } \underline{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{t-1} \\ a_0 \end{bmatrix} := \begin{bmatrix} r_0 - r_1 \\ r_0 - r_2 \\ \vdots \\ r_0 - r_{t-1} \\ 0 \end{bmatrix}.$$

(Notice that we shifted in the indices of vector $\underline{a}$ and the last coordinate a_0 is always zero.) In short, $c_j = r_j - r_{j+1}$ and $a_j = r_0 - r_j$. Let us write the relationships in matrix form. $\underline{c} = C\underline{r}$, $\underline{a} = B\underline{c}$, and $\underline{a} = (BC)\underline{a} =: A\underline{r}$, where

$$C := \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ -1 & & & 1 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \dots & 1 \end{bmatrix} \text{ and}$$

$$A := BC = \begin{bmatrix} 1 & -1 & & \\ \vdots & & \ddots & \\ 1 & & & -1 \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Matrix B has an inverse, $B^{-1} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}$, and thus we can write $\underline{c} =$

$B^{-1}\underline{a}$. The determinant of matrix C, however, is zero. Still, it possesses a right inverse in the following sense:

Let $D := \begin{bmatrix} 1 & \dots & \dots & 1 \\ -1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -1 & \dots & -1 & 1 \end{bmatrix}$, let f be any $\mathbf{C}^t \to \mathbf{C}$ function and $\underline{1}$ be a column vector with all coordinates 1.

If $\underline{r}' := f(\underline{c}) \cdot \underline{1} + \frac{1}{2}D\underline{c}$, then $C\underline{r}' = f(\underline{c}) \cdot \underline{0} + I\underline{c} = \underline{c} = C\underline{r}$.

Let us introduce two quadratic forms:

$$Q_1(\underline{x}) := \sum_{j=0}^{t-1}(x_j^2 - x_j x_{j+1}) \tag{21}$$

and

$$Q_2(\underline{x}) := \frac{1}{2}\sum_{j=0}^{t-1} x_j^2, \text{ where } \underline{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{t-1} \end{bmatrix}. \tag{22}$$

For example, if $t = 4$, $Q_1(\underline{x})$ equals $x_0^2 - x_0x_1 + x_1^2 - x_1x_2 + x_2^2 - x_2x_3 + x_3^2 - x_3x_0$ for a general $\underline{x}$. But Q_1 takes a different form when applied to the specially indexed vector $\underline{a}$:

$$Q_1(\underline{a}) = a_1^2 - a_1a_2 + a_2^2 - a_2a_3 + a_3^2, \text{ since } a_0 = 0.$$

LEMMA 1. *Let g be an arbitrary number. The following identities hold:*

$$Q_1(\underline{r} + g \cdot \underline{1}) = Q_1(\underline{r}), \tag{23}$$

$$Q_2(\underline{c}) = Q_2(C\underline{r}) = Q_1(\underline{r}), \tag{24}$$

$$Q_1(\underline{a}) = Q_1(A\underline{r}) = Q_1(\underline{r}), \tag{25}$$

$$-\sum_{j=0}^{t-1} jc_j = -\sum_{j=0}^{t-1} r_j + tr_0 = \sum_{j=0}^{t-1} a_j, \tag{26}$$

$$t \cdot Q_2(\underline{c}) - L(\underline{c}) = \frac{1}{t} \cdot Q_2(\underline{c}^*) - L^* + g\sum_{j=0}^{t-1} c_j, \tag{27}$$

$$where\ L^* := \frac{t^2-1}{24} + \frac{1}{2}(g - \frac{t-1}{2})^2 = \frac{t^2+2}{24} + \frac{1}{2}(g - \frac{t-2}{2})(g - \frac{t}{2}), \tag{28}$$

$$L(\underline{c}) := \sum_{j=0}^{t-1} jc_j, \tag{28'}$$

and

$$c_j^* := tc_j + j - g. \tag{29}$$

IN LIEU OF PROOF: The identities above are easy to check. We found (25) by writing $Q_1(\underline{r}) = \frac{1}{4}(2r_0 - r_1 - r_{t-1})^2 + \frac{1}{2}(r_1 - r_2)^2 + \ldots + \frac{1}{2}(r_{t-2} - r_{t-1})^2 + \frac{1}{4}(r_{t-1} - r_1)^2$, which made us introduce the coordinate system $\underline{a}$ with $a_0 = 0$ as defined above. ∎

§2.5. More Identities

Let us remind the reader of the notations (28') and (22) of the last section, $L(\underline{c}) = \sum_{j=0}^{t-1} jc_j$ and $Q_2(\underline{c}) = \frac{1}{2}\sum_{j=0}^{t-1} c_j^2$. If $\sum_{j=0}^{t-1} c_j = 0$ (and certainly this is true for the numbers c_j defined by $r_j - r_{j+1}$ there) $Q_2(\underline{c})$ is equal to $\sum_{j=1}^{t-1} c_j^2 + \sum_{0<j<k} c_j c_k$.

As in §2, let ω denote $e^{\frac{2\pi i}{t}}$. The identities

$$(z;q)_\infty = (z, zq, \ldots, zq^{t-1}; q^t)_\infty \tag{30}$$

and

$$(z^t;q^t)_\infty = (z, z\omega, \ldots, z\omega^{t-1}; q)_\infty \tag{30'}$$

will come handy in proving the following lemma.

LEMMA 2. *Let* $\sum_{\underline{c}\geq -k}$ *denote summation over*

$$\{\underline{c}=[c_0,c_1,\ldots,c_{t-1}]:\sum_{j=0}^{t-1}c_j=0 \text{ and } c_0,c_1,\ldots,c_{t-1}\geq -k\}.$$

We have

$$\frac{1}{(q)_k}=\sum_{\underline{c}\geq -k}\frac{q^{t\cdot Q_2(\underline{c})+L(\underline{c})}}{\prod_j (q^t;q^t)_{c_j+k}} \tag{31}$$

and, if t is an odd number,

$$\frac{1}{(q^t;q^t)_k}=\sum_{\underline{c}\geq -k}\frac{q^{Q_2(\underline{c})}\omega^{L(\underline{c})}}{\prod_j (q)_{c_j+k}} \tag{31'}$$

PROOF: We have to use

$$(-x)_\infty=\sum_{k=0}^{\infty}\frac{q^{\binom{k}{2}}}{(q)_k}x^k. \tag{32}$$

If we substitute x^t for x, then q^t for q, we get

$$(-x^t;q^t)_\infty=\sum_{k=0}^{\infty}\frac{q^{t\binom{k}{2}}}{(q^t;q^t)_k}x^{tk}. \tag{32'}$$

Apply (32) to both sides of (30) to obtain

$$\sum_{k=0}^{\infty}\frac{q^{\binom{k}{2}}}{(q)_k}x^k=(-x)_\infty=\sum_{k_0,\ldots,k_{t-1}=0}^{\infty}\frac{q^{t\sum\binom{k_j}{2}}}{\prod_j (q^t;q^t)_{k_j}}q^{\sum jk_j}x^{\sum k_j}. \tag{33}$$

Similarly, from (32') and (30'), provided $(-x)^t=-x^t$, i.e. t is odd, we gain

$$\sum_{k=0}^{\infty}\frac{q^{t\binom{k}{2}}}{(q^t;q^t)_k}x^{tk}=(-x^t;q^t)_\infty=\sum_{k_0,\ldots,k_{t-1}=0}^{\infty}\frac{q^{\sum\binom{k_j}{2}}}{\prod_j (q)_{k_j}}\omega^{\sum jk_j}x^{\sum k_j}. \tag{33'}$$

Extract the coefficients of x^{tk} from both sides to get

$$\frac{q^{\binom{tk}{2}}}{(q)_{tk}} = \sum_{\sum k_j = tk, k_j \geq 0} \frac{q^{t\sum\binom{k_j}{2}}}{\prod_j (q^t;q^t)_{k_j}} q^{\sum jk_j} \tag{34}$$

and

$$\frac{q^{t\binom{k}{2}}}{(q^t;q^t)_k} = \sum_{\sum k_j = tk, k_j \geq 0} \frac{q^{\sum\binom{k_j}{2}}}{\prod_j (q)_{k_j}} \omega^{\sum jk_j}. \tag{34'}$$

Introduce $c_j := k_j - k$. Remark that

$$\sum_{j=0}^{t-1} k_j = tk \text{ iff } \sum_{j=0}^{t-1} c_j = 0 \tag{35}$$

Finally, use the identities

$$\sum_{j=0}^{t-1} \binom{c_j + k}{2} = \frac{1}{2}\sum_{j=0}^{t-1} c_j^2 + \left(k - \frac{1}{2}\right)\sum_{j=0}^{t-1} c_j + t\cdot\binom{k}{2} \tag{36}$$

and $\binom{tk}{2} = t^2\binom{k}{2} + k\sum_{j=0}^{t-1} j$ to obtain (31). If t is odd, then $\omega^{k\sum j} = 1$ and (36) give (31'). ∎

Define

$$\eta^*[t] := (q^t;q^t)_\infty \text{ and } \eta[t] := q^{\frac{t}{24}}\eta^*[t] \tag{37}$$

Recall that the usual η-function is defined by

$$\eta(x) := q_x^{\frac{1}{24}}(q_x)_\infty, \text{ where } q_x := e^{2\pi i x}.$$

Therefore, by substituting $q = e^{2\pi i\tau}$ in (37) we obtain $\eta[t] = \eta(t\cdot\tau)$.

If we take $k \to \infty$ in Lemma 2, we obtain

COROLLARY 1.

$$\frac{\eta^*[t]^t}{\eta^*[1]} = \sum_{\underline{c}} q^{t\cdot Q_2(\underline{c})+L(\underline{c})} \tag{38}$$

and, for odd t,

$$\frac{\eta^*[1]^t}{\eta^*[t]} = \sum_{\underline{c}} q^{Q_2(\underline{c})}\omega^{L(\underline{c})}, \tag{38'}$$

where the summation $\sum_{\underline{c}}$ *extends for all vectors* $\underline{c} = [c_0, c_1, \ldots, c_{t-1}]$ *such that* $\sum_{j=0}^{t-1} c_j = 0$.

§2.6. Various Expressions for the Number of t-cores

We can prove (38) combinatorially too. Let P and P_t^* denote the sets of all partitions and all t-core partitions respectively. The following two lemmas are quoted from [Gar90]:

LEMMA 3. *There is a bijection between*

$$\pi \in P \text{ and } [\pi_0, \ldots, \pi_{t-1}, \pi^*] \in P \times \ldots \times P \times P_t^*, \tag{39}$$

which satisfies

$$|\pi| = t \cdot \sum_{j=0}^{t-1} |\pi_j| + |\pi^*|. \tag{40}$$

Equation (40) yields

$$\sum_{n=0}^{\infty} p(n)q^n = \frac{1}{(q^t;q^t)_\infty^t} \cdot \left(\sum_{n=0}^{\infty} p_t^*(n)q^n\right) \tag{41}$$

LEMMA 4. *There is a second bijection between*

$$\pi^* \in P_t^* \text{ and } \underline{c} \in \{\underline{c} = [c_0, c_1, \ldots, c_{t-1}] \in Z^t : \sum_{j=0}^{t-1} c_j = 0\} \tag{42}$$

which satisfies

$$|\pi^*| = t \cdot Q_2(\underline{c}) + L(\underline{c}) \tag{43}$$

Equation (43) yields

$$\sum_{n=0}^{\infty} p_t^*(n) q^n = \sum_{\underline{c}} q^{t \cdot Q_2(\underline{c}) + L(\underline{c})}, \tag{44}$$

where $Q_2(\underline{c})$, $L(\underline{c})$ and $\sum_{\underline{c}}$ were defined in (22), (28') and Corollary 1, respectively.

A SECOND PROOF OF (38) OF COROLLARY 1:

$$\sum_{\underline{c}} q^{t \cdot Q_2(\underline{c}) + L(\underline{c})} = \sum_{n=0}^{\infty} p_t^*(n) q^n = (q^t; q^t)_\infty^t \cdot \sum_{n=0}^{\infty} p(n) q^n = \frac{(q^t; q^t)_\infty^t}{(q)_\infty}. \quad \blacksquare$$

NOTE.

Let $p_t^*(n)$ denote the number of t-cores π satisfying $|\pi| = n$. From the table above we can see that $p_3^*(3) = p_3^*(7) = 0$, for instance. $p_2^*(n)$ can easily be zero too: $p_2^*(n)$ equals 1 or 0 as n equals a triangular number $\binom{m}{2}$ or not. On the other hand, Ken Ono [Ono93] proved in his dissertation that $p_{2k}(n) > 0$ ($k = 2, 3, \ldots$ $n = 1, 2, \ldots$). It is conjectured that $p_{2k+1}(n) > 0$ for every $k = 2, 3, \ldots$.

We will give several ways to calculate $p_t^*(n)$ in the next theorem.

THEOREM 2. Let $\#\{\ldots\}$ denote the number of elements in set $\{\ldots\}$. *The following numbers are all equal to $p_t^*(n)$, i.e. the number of t-core partitions of n:*

$$\#\{\underline{r} \in \mathbf{N}^t : \sum_{j=0}^{t-1} r_j = n, \; Q_1(\underline{r}) = r_0\}, \tag{45}$$

$$\#\{\underline{c} \in \mathbf{Z}^t : t \cdot Q_2(\underline{c}) + \sum_{j=0}^{t-1} j c_j = n, \; \sum_{j=0}^{t-1} c_j = 0\}, \tag{46}$$

$$\frac{\eta^*[t]^t}{\eta^*[1]} // q^n, \tag{47}$$

$$\frac{\eta[t]^t}{\eta[1]}//q^{n+\frac{t^2-1}{24}}, \tag{48}$$

$$\#\{\underline{c}^* \in \mathbf{Z}^t : \frac{1}{t}Q_2(\underline{c}^*) = n + L^*,\ \sum_{j=0}^{t-1} c_j^* = \frac{t(t-1)}{2} - t\cdot g$$
$$c_j^* \equiv j - g \pmod{t}\} \tag{49}$$

$$\#\{\underline{c}^* \in \mathbf{Z}^t : \frac{1}{t}Q_2(\underline{c}^*) = n + \frac{t^2-1}{24},\ \sum_{j=0}^{t-1} c_j^* = 0$$
$$c_j^* \equiv j - \frac{t-1}{2} \pmod{t}\},\ \textit{if } t \textit{ is odd}\ , \tag{49a}$$

$$\#\{\underline{c}^* \in \mathbf{Z}^t : \frac{1}{t/2}Q_2(\underline{c}^*) = 2n + \frac{t^2+2}{12},\ \sum_{j=0}^{t-1} c_j^* = \frac{t}{2}$$
$$c_j^* \equiv j - \frac{t-2}{2} \pmod{t}\},\ \textit{if } t \textit{ is even}\ , \tag{49a}$$

$$\#\{\underline{a} \in \mathbf{Z}^{t-1} : t\cdot Q_1(\begin{bmatrix}\underline{a}\\0\end{bmatrix}) - \sum_{j=1}^{t-1} a_j = n\}. \tag{50}$$

NOTE.

Equations (45)-(48) are known, they appear in [Gar90], where $\underline{c}$ is denoted by $\underline{n}$, (49a) is mentioned without the explicit relationship (27) in [Kly82], (50) seems to be new and this is the coordinate system we prefer to use.

PROOF: The expression in (46) is equal to $p_t^*(n)$ by the bijection of Lemma 4. (46) $\Longleftrightarrow$ (45): by (24) and (26), we have $n = tQ_2(\underline{c}) - L(\underline{c}) = tQ_1(\underline{r}) + \sum_{j=0}^{t-1} r_j - tr_0 = \sum_{j=0}^{t-1} r_j$, so we see that $p_t(n) \leq (45)$. As we mentioned in §4, matrix C has a right inverse. The explicit relationship

$$\underline{r} = Q_2(\underline{c})\cdot\underline{1} + \frac{1}{2}D\underline{c} \tag{51}$$

(see [Gar90, (4.1)] shows that (45) $\leq p_t(n)$.

(46) $\Longleftrightarrow$ (49) comes from (27).

(46) $\Longleftrightarrow$ (50), since $\det B = 1 \neq 0$ and thus B gives a bijection between $\underline{c}$ and $\underline{a}$. For the linear part we can use (26).

(46) $\Longleftrightarrow$ (47) is (38) of Corollary 1.

(47) $\Longleftrightarrow$ (48) comes from definition (30). ∎

COROLLARY 2. *Let $K \geq 1$ and $e(k) \geq 1$ for $1 \leq k \leq K$. Then*

$$\#\{\underline{\pi} := [\pi(1), \pi(2), \ldots, \pi(K)] : \pi(k) \textit{ is a } t(k)\textit{-core and } \sum_{k=1}^{K} e(k) \cdot \#\pi(k) = n\}$$

$$= \prod_{k=1}^{K} \frac{\eta^*[e(k) \cdot t(k)]^{t(k)}}{\eta^*[e(k)]} // q^n = \prod_{k=1}^{K} \frac{\eta[e(k) \cdot t(k)]^{t(k)}}{\eta[e(k)]} // q^{n + \sum_{k=1}^{K} e(k) \cdot \frac{t(k)^2 - 1}{24}}$$

$$= \#\{\underline{a} \in \mathbf{Z}^d : \sum_{k=1}^{K} e(k) t(k) Q_1(\begin{bmatrix} \underline{a}(k) \\ 0 \end{bmatrix}) - \underline{e} \cdot \underline{a} = n\},$$

where $d = \sum_{k=1}^{K}(t(k) - 1)$, $\underline{a} := [\underline{a}(1), \ldots, \underline{a}(K)]^c$, $\underline{e} := [\underline{e}(1), \ldots, \underline{e}(K)]$ and $\underline{e}(k) = e(k) \cdot \underline{1} \in \mathbf{Z}^{t(k)-1}$.

COROLLARY 3.

If, in the process mentioned in §3 we have to remove $t \cdot k$ lattice points from π to get to the core , we have the following identity for $r_j = r_j(\pi)$:

$$r_0 - k = \sum_{j=0}^{t-1} (r_j^2 - r_j r_{j+1}) \quad \textit{where } r_t := r_0.$$

PROOF: For each j, r_j decreases by the same number when we remove a rimhook. But the right hand side of the claim does not change, by (23) of Lemma 1. Therefore $r_0 - \sum_{j=0}^{t-1}(r_j^2 - r_j r_{j+1})$ has decreased by k to 0 when we reach the t-core (see (45)). ∎

Let r_j be defined as in §1, i.e. the number of dots colored j in the t-residue diagram. As an illustration of the three characterizations of t-cores via $\underline{r}, \underline{c}, \underline{a}$ we include the following table.

TABLE 2-1. 3-CORE PARTITIONS OF INTEGERS ≤ 20

size	partition	a_1, a_2	r_0, r_1, r_2	c_0, c_1, c_2
1	=1	1,1	1,0,0	1,0,-1
2	=2	0,1	1,1,0	0,1,-1
	=1+1	1,0	1,0,1	1,-1,1
4	=3+1	0,-1	1,1,2	0,-1,1
	=2+1+1	-1,0	1,2,1	-1,1,0
5	=3+1+1	-1,-1	1,2,2	-1,0,1
6	=4+2	2,1	3,1,2	2,-1,-1
	=2+2+1+1	1,2	3,2,1	1,1,-2
8	=4+2+1+1	2,2	4,2,2	2,0,-2
9	=5+3+1	-1,1	3,4,2	-1,2,-1
	=3+2+2+1+1	1,-1	3,2,4	1,-2,1
10	=5+3+1+1	0,2	4,4,2	0,2,-2
	=4+2+2+1+1	2,0	4,2,4	2,-2,0
12	=6+4+2	-1,-2	3,4,5	-1,-1,2
	=3+3+2+2+1+1	-2,-1	3,5,4	-2,1,1
14	=6+4+2+1+1	0,-2	4,4,6	0,-2,2
	=5+3+2+2+1+1	-2,0	4,6,4	-2,2,0
16	=7+5+3+1	3,2	7,4,5	3,-1,-2
	=4+3+3+2+2+1+1	2,3	7,5,4	2,1,-3
	=6+4+2+2+1+1	-2,-2	4,6,6	-2,0,-2
17	=7+5+3+1+1	3,1	7,4,6	3,-2,-1
	=5+3+3+2+2+1+1	1,3	7,6,4	1,2,-3
20	=4+4+3+3+2+2+1+1	2,-1	7,5,8	2,-3,1
	=8+6+4+2	-1,2	7,8,5	-1,3,-2

§2.7. A New Proof of the Cranks for $p_t^*(tn-(t^2-1)/24)$

We will use coordinate system $\underline{a}$ in the rest of the chapter. It leads to more elegant proofs than $\underline{c}$ used in [Gar90] (which is denoted by $\underline{n}$ there). Let us remind the reader that although the vector $\underline{a} = \underline{a}(\pi)$ was introduced as a t-dimensional column vector with $a_j = r_0 - r_j$ in §4, but we can cut off its last, $a_0 = 0$ coordinate. Thus $\underline{a}$ denotes a $(t-1)$- dimensional vector from now on. (It denoted the same, truncated vector in §6 too.)

Row vector $\underline{f}$ is called a *crank* for the t-core solutions of $|x| = n$, if for each k, $0 \le k \le t-1$,

$$\#P_t^*\left(n, \underline{f}\cdot\underline{a} \equiv k \pmod t\right) = \frac{\#P_t^*(n)}{n},$$

where, by (50),

$$\#P_t^*(n) = \{\underline{a} \in \mathbf{Z}^{t-1} : n = Q_1(\begin{bmatrix}a\\0\end{bmatrix}) - \sum_{j=0}^{t-1} a_j\}. \tag{52}$$

Remark that the row vector $\underline{e}$ of Corollary 2 consists of all coordinates 1, i.e. $\underline{e} = \underline{1}$ for the problem $|x| = n$. Actually, we would like to find cranks for

$$|x| = tn - \frac{t^2-1}{24} = tn - L,$$

where

$$L := 1, 2, 5 \text{ (for } t = 5, 7, 11).$$

and

$$L(\underline{a}) := \underline{e}\cdot\underline{a} = a_1 + a_2 + \ldots + a_{t-1} \equiv L \pmod t. \tag{53}$$

THEOREM 3. *Let t=5,7,11. The vectors*

$\underline{f}_5 = [3,1,-1,-3]$, $\underline{f}_7 = [2,1,1,-1,-1,-2]$, $\underline{f}_{11} = [-5,4,2,2,1,-1,-2,-2,-4,5]$ *are cranks for the problem* $|x| = tn - \frac{t^2-1}{24}$.

NOTE.

The cranks are given in terms of the $\underline{c}$ coordinate system in [Gar90], we just had to calculate them in terms of our $\underline{a}$. The proof below, however, is new.

PROOF: We will construct affine transformations $T_t : \mathbf{Z}^{t-1} \to \mathbf{Z}^{t-1}$ which give bijections $P_t^*\left(tn - L, \underline{f} \cdot \underline{a} \equiv k \pmod{t}\right) \leftrightarrow P_t^*\left(tn - L, \underline{f} \cdot \underline{a} \equiv k+1 \pmod{t}\right)$. We break down the proof into several steps.

Step 1. Encode the partitions of the sets $P_t^*(t-L)$, (which is the same as $P_t(t-L)$, since $t - L < t$) into $\underline{a}$. The vectors $\underline{a}$ appear in the columns of the following table (the last coordinates a_0 are always zero, and are omitted):

TABLE 2-2. THE PARTITIONS OF $t-L$ IN THE COORDINATE SYSTEM $\underline{a}$.

$P_5(4) =$
0 0 1 0 1
0 0 2 1 0
0 1 2 0 0
1 0 1 0 0
$\underline{f}_5 \cdot P_5(4) =$
2 4 0 1 3

$P_7(5)=$
0 0 1 0 1 0 1
0 0 1 0 2 1 1
0 0 2 1 2 1 0
0 1 2 1 2 0 0
1 1 2 0 1 0 0
1 0 1 0 1 0 0
$\underline{f}_7 \cdot P_7(5) =$
4 5 6 0 1 2 3

$P_{11}(6) =$
0 0 0 1 0 1 1 1 0 0 1
0 0 0 1 1 1 2 2 0 1 1
0 0 0 1 2 2 2 2 1 1 1
0 0 1 2 2 2 2 2 1 1 1
0 1 1 2 2 2 2 2 1 1 1
1 1 1 2 2 2 2 2 1 1 0
1 1 1 2 2 2 2 2 1 0 0
1 1 1 2 2 2 2 1 0 0 0
1 1 0 2 2 1 1 1 0 0 0
1 0 0 1 1 1 0 1 0 0 0
$\underline{f}_{11} \cdot P_{11}(6) =$
7 3 9 5 1 0 A 6 2 8 4,
where A=10

The number of columns, $p_t(t-L)$ equals t and multiplication by $\underline{f}_t$ above shows that the (hoped) cranks of the partitions of $P_t(t-L)$ belong to different congruence classes $\pmod{t}$.

Step 2. Any affine transformation T can be written in a matrix form (in this section $d = t - 1$):

$$T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) := \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \dots & \underline{x}_d & \underline{b} \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{a} \\ 1 \end{bmatrix} := \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1,d} & b_1 \\ x_{21} & x_{22} & \dots & x_{2,d} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{d,1} & x_{d,2} & \dots & x_{d,d} & b_d \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \\ 1 \end{bmatrix}$$

This has the advantage that composition of affine transformations can be calculated by matrix multiplication. Let us emphasize that here we substituted the last coordinates $a_0 = 0$ with $a_0 = 1$, which means we will work with the *dilated* vectors $\underline{a}$! Let $^*\underline{v}$ denote the significant part of $\underline{v}$, i.e. write

$$T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) = \begin{bmatrix} ^*T([\begin{smallmatrix} \underline{a} \\ 1 \end{smallmatrix}]) \\ 1 \end{bmatrix}.$$

Define the affine transformations T_t by two requirements:

$$\text{if } \underline{a} \in P_t(t-L), \text{ then } ^*T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) \in P_t(t-L), \tag{54}$$

$$[\underline{f}, 0] \cdot T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) = \underline{f} \cdot \underline{a} + 1. \tag{55}$$

Matrix T_t permutes the columns of the encoded $P_t(t-L)$ in the order determined by (55), therefored the desired T_t must be defined by

$$T_t := [\text{permuted } P_t(t-L)] \cdot [P_t(t-L)]^{-1}.$$

The matrices obtained are given below:

TABLE 2-3. THE MATRICES OF THE AFFINE TRANSFORMATIONS T_t.

$5 \cdot T_5$				
-3	-3	2	2	3
-3	-3	7	-3	3
0	-5	5	-5	5
- 4	1	1	-4	4
0	0	0	0	5

$7 \cdot T_7$						
-6	1	1	1	1	-6	5
-7	0	0	7	-7	0	7
-3	-3	-3	11	-3	-3	6
-1	-8	-1	6	-1	-1	9
-1	-1	-8	6	-1	-1	9
4	-3	-3	4	-3	-3	6
0	0	0	0	0	0	7

$11 \cdot T_{11}$										
3	3	-8	3	3	-8	3	3	3	-8	7
9	-2	-13	9	-2	-2	-2	9	-2	-2	10
7	-4	-4	7	-4	-4	-4	18	-4	-4	9
8	-3	-3	8	-3	-3	-14	19	-3	-3	15
12	-10	1	12	-10	1	-10	12	1	-10	17
8	-3	-3	19	-14	-3	-3	8	-3	-3	15
7	-4	-4	18	-4	-4	-4	7	-4	-4	9
9	-2	-2	9	-2	-2	-2	9	-13	-2	10
14	-8	3	3	3	-8	3	3	-8	3	7
11	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	11

The first salient feature of the matrices T_t above is that they are $= \frac{1}{t}\cdot$ (integer matrix). We will show that these transformations act analogously to (54) and (55) on $P_t^*(tn - L)$, for every $n = 1, 2, 3, \ldots$

Step 3. First of all, they satisfy $T^t = I$. This was verified by hand and Matlab.

Step 4. They preserve the size of the t-core:

$$Q_1\left(\begin{bmatrix} {}^*T(\left[\begin{smallmatrix} a \\ 1 \end{smallmatrix}\right]) \\ 0 \end{bmatrix}\right) - L\left(\begin{bmatrix} {}^*T(\left[\begin{smallmatrix} a \\ 1 \end{smallmatrix}\right]) \\ 0 \end{bmatrix}\right) = Q_1(\begin{bmatrix} a \\ 0 \end{bmatrix}) - L(\begin{bmatrix} a \\ 0 \end{bmatrix}) \tag{56}$$

The last coordinate of vector $T(\left[\begin{smallmatrix} a \\ 1 \end{smallmatrix}\right])$ is always 1, and should be replaced by 0 before Q_1 or L are applied on the left hand side of (56). Condition (56) is a very strong one, and this is the easiest place for the crank property to break down. We used Mathematica to verify (56).

Step 5. Check that the if $\underline{a} \in \mathbf{Z}^d$ then $T(\underline{a}) \in \mathbf{Z}^d$ too. Well, this follows from the beautiful structure of these matrices. All entries of row j (except the last, dilational

one), are congruent to the same number $\xi_j \pmod t$. In other words, if we write $t \cdot T(\left[\begin{smallmatrix} \underline{a} \\ 1 \end{smallmatrix}\right]) = X\underline{a} + \underline{b}$, $\underline{e} = [1, \ldots, 1] \in \mathbf{Z}^{t-1}$ and $\underline{\xi} := \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{t-1} \end{bmatrix}$, then we have

$$X \equiv \underline{\xi} \cdot \underline{e} \pmod t. \tag{57}$$

In addition, observe that

$$L\underline{\xi} + \underline{b} \equiv 0 \pmod t. \tag{58}$$

Therefore

$$t \cdot T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) = X\underline{a} + \underline{b} \equiv \underline{\xi}(\underline{e} \cdot \underline{a}) + \underline{b} \equiv \underline{\xi}L + \underline{b} \equiv \underline{0} \pmod t.$$

Step 6. There are integers α and β such that

$$[\underline{f}, 0]T - [\underline{f}, 0] \equiv \alpha \cdot [\underline{e}, 0] + \beta \cdot [\underline{0}, 1] \tag{59}$$

and

$$L\alpha + \beta \equiv 1 \pmod t. \tag{60}$$

The values of α and β can be found in Table at the end of the proof of this theorem.

Therefore for any $\underline{a} \in \mathbf{Z}^{t-1}$ we have $\underline{f} \cdot^{*} T(\left[\begin{smallmatrix} \underline{a} \\ 1 \end{smallmatrix}\right]) - \underline{f} \cdot \underline{a} \equiv \alpha \underline{e}\underline{a} + \beta \equiv 1 \pmod t$, proving that T_t maps $P_t^*(n, \underline{f} \cdot \underline{a} \equiv k \pmod t)$ into $P_t^*(n, \underline{f} \cdot \underline{a} \equiv k + 1 \pmod t)$

Step 7. The determinant of T_t is 1 for each $t = 5, 7, 11$, therefore T_t establishes a bijection. This finishes the proof that for every k the sets $P_t^*(n, \underline{f}_t \cdot \underline{a} \equiv k \pmod t)$ contain the same number of elements, where $\underline{f}_t$ are the vectors given in the statement of Theorem 3.

With this, we finished the proof of Theorem 3. ∎

TABLE 2-4. SOME PARAMETERS OF T_t.

t	L	α	β	ξ'	h
5	1	2	-1	-2	1
7	2	1	-1	1	1
11	5	-4	-1	4	2

ADDITIONAL PROPERTIES OF THE MATRICES T_t.

The entries of row j of matrix X are all congruent to

$$1 - \left(\xi' + (j-1)h\right)^2 \pmod{t}, \tag{61}$$

where the values of ξ' and h are given in Table 4. above.

We can state a property like (57) for any power of T:

For each k we have

$$t \cdot T^k \equiv \begin{bmatrix} \underline{\xi}(k)\underline{e} & \underline{b}(k) \\ \underline{0} & 1 \end{bmatrix} \tag{62}$$

The sets $\{\xi_j(k) : j = 1, 2, \ldots, t-1\}$ are permutations of the roots of the following polynomials of degree $t-1$:

$$R(x) = x(x-1) \cdot \prod_{j=2}^{\frac{t-1}{2}} \left(x - 1 + j^2\right)^2 = \prod_{j=1}^{t-1} \left(x - 1 + (j-1)^2\right). \tag{63}$$

This shows a one-to-one correspondence between the powers of roots of unity, ω^j and $1-(j-1)^2$.

Now concatenate the columns of ξ_j to get the matrix Ξ_t. For example,

$$\Xi_7 := [\underline{\xi}(1), \underline{\xi}(1), \underline{\xi}(2), \underline{\xi}(3), \underline{\xi}(4), \underline{\xi}(5), \underline{\xi}(6)] = \begin{bmatrix} 1 & 0 & -3 & -1 & -1 & -3 \\ 0 & -1 & -3 & 1 & -3 & -1 \\ -3 & -3 & 0 & -1 & 1 & -1 \\ -1 & 1 & -1 & 0 & -3 & -3 \\ -1 & -3 & 1 & -3 & -1 & 0 \\ -3 & -1 & -1 & -3 & 0 & 1 \end{bmatrix}$$

We can observe that

$$\det[\Xi_t] = -t^{t-3}, \tag{64}$$

provided we choose $\xi_j(k) \in [-\frac{t-1}{2}, \frac{t-1}{2}]$.

Let us make a final observation: the matrices $M = 11 \cdot T_{11}, 7 \cdot T_7^{-2}$ and $5 \cdot T_5^2$ have remarkably similar structures.

TABLE 2-5. MATRICES WITH SIMILAR STRUCTURE.

$5 \cdot T_5^2$				
2	2	-3	-3	3
6	-4	1	-4	4
7	-3	-3	2	3
5	0	0	0	0
0	0	0	0	5

$7 \cdot T_7^{-2}$						
-1	6	-8	6	-1	-1	2
4	4	-10	4	-3	4	6
1	8	-6	1	-6	8	5
4	4	-3	4	-10	4	6
6	-1	-1	6	-8	6	2
7	0	0	0	0	0	0
0	0	0	0	0	0	7

$11 \cdot T_{11}$										
3	3	-8	3	3	-8	3	3	3	-8	7
9	-2	-13	9	-2	-2	-2	9	-2	-2	10
7	-4	-4	7	-4	-4	-4	18	-4	-4	9
8	-3	-3	8	-3	-3	-14	19	-3	-3	15
12	-10	1	12	-10	1	-10	12	1	-10	17
8	-3	-3	19	-14	-3	-3	8	-3	-3	15
7	-4	-4	18	-4	-4	-4	7	-4	-4	9
9	-2	-2	9	-2	-2	-2	9	-13	-2	10
14	-8	3	3	3	-8	3	3	-8	3	7
11	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	11

In all three matrices M above, we have $M(t-1, 1) = t$, $M(t-1, 2..t) = \underline{0}$, $M(1..t-2, 2..t-1)$ are centrally symmetric submatrices with $M(\frac{t-1}{2}, \frac{t+1}{2}) = 1$ in their center. In addition, $M(\frac{t-1}{2}, \frac{t-1}{2}) = -(t-1)$. Of course, similar facts can be said about $M = 11 \cdot T_{11}^{-1}, 7 \cdot T_7^2$ and $5 \cdot T_5^{-2}$ too.

§2.8. An Identity of Nathan Fine's

In (37) of §5 we already introduced $\eta[t] := q^{\frac{t}{24}}(q^t;q^t)_\infty$, where $q = e^{2\pi i\tau}$.

In N. Fine's book [Fin88, page 86] we find the following identity:

$$\frac{\eta[1\cdot 3]^3}{\eta[1]}\cdot\frac{\eta[2\cdot 3]^3}{\eta[2]}//q^N = \sum_{d|N} d\cdot f(d), \quad \text{where } f(d) = \begin{cases} 0, & \text{if } d \equiv 0, \pm 2 \\ \frac{2}{3}, & \text{if } d \equiv 3 \\ 1, & \text{if } d \equiv \pm 1 \pmod 6. \end{cases} \tag{65}$$

Take first N = prime, then N = prime power, finally a general N to obtain the following lemma.

LEMMA 5.

If $N = 2^\kappa\cdot 3^\lambda\cdot p_1^{\mu_1}\cdot\ldots\cdot p_s^{\mu_s}$ *then*

$$\frac{\eta[1\cdot 3]^3}{\eta[1]}\cdot\frac{\eta[2\cdot 3]^3}{\eta[2]}//q^N = 3^\lambda\cdot\frac{p_1^{\mu_1+1}-1}{p_1-1}\cdot\ldots\cdot\frac{p_s^{\mu_s+1}-1}{p_s-1}. \tag{66}$$

Let us specialize Corollary 2. in §6. Here $K = 2$, $t(1) = t(2) = 3$, $d = (3-1)+(3-1) = 4$ and

$$\underline{e} = [1,1,2,2]. \tag{67}$$

Apply Lemma 5 with $N = n + 1\cdot\frac{3^2-1}{24} + 2\cdot\frac{3^2-1}{24} = n+1$ to obtain the following theorem.

THEOREM 4.

If $n = 2^\kappa\cdot 3^\lambda\cdot p_1^{\mu_1}\cdot\ldots\cdot p_s^{\mu_s} - 1$ *then*

$$\begin{aligned} s^*(n) &= \#\{[\pi(1),\pi(2)] : \textit{ both are 3-cores and } |\pi(1)| + 2\cdot|\pi(2)| = n\} \\ &= 3^\lambda\cdot\frac{p_1^{\mu_1+1}-1}{p_1-1}\cdot\ldots\cdot\frac{p_s^{\mu_s+1}-1}{p_s-1} \\ &= \#\{\underline{a}\in\mathbb{Z}^{2+2} : 3(a_1^2+a_2^2-a_1a_2)+6(a_3^2+a_4^2-a_3a_4)-(a_1+a_2)-2(a_3+a_4) = n\}. \end{aligned}$$

We would like point out two consequences of Theorem 4:

$$s^*(3n+2) = 3 \cdot s^*(n) \tag{68}$$

and

$$s^*(2n+1) = s^*(n). \tag{69}$$

§2.9. Two Cranks for $s^*(3n+2)$

First we have two identify our crank candidates among the vectors of $\mathbf{Z}^{2\cdot 2}$ (mod 3). Each of the four coordinates can have values $\{-1, 0, 1\}$. That gives 3^4 choices for $\underline{f}$. We can cut down the size of the list of candidates by regarding $|\pi_1|+2\cdot|\pi_2| = 3\cdot 6+2 = 3\cdot 7-1$. This equation has $3\cdot\frac{7^2-1}{7-1} = 24$ solutions of 3-core pairs symbolized by $\underline{a} = [a_1, a_2, a_3, a_4]$. Only six out of the 81 vectors classify these solutions into three equinumerous classes. But two of these six vectors, $\underline{f} = [-1, 1, 0, 0]$ and $[1, -1, 0, 0]$ do not classify the 18 solutions of $|\pi_1| + 2\cdot|\pi_2| = 3\cdot 4 + 2 = 3\cdot 5 - 1$ evenly, $\underline{f}\cdot\underline{a} \equiv 0 \pmod 3$ does not occur at all, while $\equiv \pm 1 \pmod 3$ occur nine times each.

So we are left with four candidates, but with only two linearly independent ones:

$$\pm[-1, 1, -1, 1] \text{ and } \pm[-1, 1, 1, -1].$$

We will prove that they are cranks indeed. But first let us show that the Andrews-Garvan crank on $S(3n+2)$ discussed in §2 is not a crank on its restriction $S^*(3n+2)$, where

$$S^*(n) := \{[\pi(1), \pi(2)] \in P_3^* \times P_3^* : |\pi(1)| + 2|\pi(2)| = n\}$$

Let us calculate the cranks of the colored 3-core solutions of $|\pi_1| + 2 \cdot |\pi_2| = 8$ (the reader might find it useful to look at Table 1).

For example, the solution $[\pi_1, \pi_2] = [2+1+1, 1+1]$ is equivalent to the colored partition $2_0 + 1 + 1 + 2_1 + 2_1$, which has two 1-colored and one 0-colored even parts, therefore its crank is $2 - 1 = 1$. On the other hand, if we calculate the scalar product of [-1,1,-1,1] with the $\underline{a}$-representation of this solution, [-1,0,1,0], the result is 0. (Multiplication of the second crank candidate [-1,1,1,-1] with [-1,0,1,0] results in -1).

Similarly, solution $[\pi_1, \pi_2] = [2+1+1, 2]$ is equivalent to the colored partition $2_0 + 1 + 1 + 4_1$, whose Andrews-Garvan crank is $1 - 1 = 0$; while the dot product of [-1,1,-1,1] by [-1,0,0,1] is congruent to $-1 \pmod 3$.

We find that the Andrews-Garvan type crank of §2 classifies the nine solutions of $|\pi_1| + 2 \cdot |\pi_2| = 8$ into non-equinumerous classes; it places (4,2,3) of them in classes (-1,0,1) (mod 3). On the other hand, both [-1,1,-1,1] and [-1,1,1,-1] classify (3,3,3) of the solutions in these classes.

THEOREM 5. *The vectors*

$$\underline{f}(1) := [-1, 1, -1, 1] \text{ and } \underline{f}(2) := [-1, 1, 1, -1]$$

are both cranks for the sets $S^*(3n+2)$.

PROOF: We will follow the steps of the proof of Theorem 3. (Historically, however, the proof of Theorem 3. is the imitation of this proof).

Step 1. The matrix of an affine transformation which proves the crank property (if it exists at all) must have $2 \cdot 2 + 1$ columns. These columns are the unknowns we are looking for. So it is not enough to encode the three partitions of $P_3(3-1)$ this

time, we have to encode another $P_3(3n+2)$ set too. To our luck, $P_3(2 \cdot 3 - 1)$ will do job, and the resulting six linear equations will give two solution quintets.

TABLE 2-6. THE PARTITIONS OF 2 AND 5 IN THE COORDINATE SYSTEM $[a_1, a_2 | a_3, a_4]$.

$P_3(2)$	$\underline{v}_1$	$\underline{v}_2$	$\underline{v}_3$	$P_3(5)$	$\underline{w}_1$	$\underline{w}_2$	$\underline{w}_3$
	1	0	0		-1	1	1
	0	1	0		-1	1	1
	0	0	1		0	1	0
	0	0	1		0	0	1
$\underline{f}(1) \cdot P_3(2)$	2	1	0	$\underline{f}(1) \cdot P_3(5)$	0	2	1
$\underline{f}(2) \cdot P_3(2)$	2	1	0	$\underline{f}(2) \cdot P_3(5)$	0	1	2

Step 2. Define the affine transformation T_3 by the requirements

$$\text{if } \underline{a} \in P_3(2), \text{ then } {}^*T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix} \in P_3(2), \tag{70'}$$

$$\text{if } \underline{a} \in P_3(5), \text{ then } {}^*T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) \in P_3(5), \tag{70''}$$

and

$$[\underline{f}, 0] \cdot T(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) \equiv [\underline{f}, 0] \cdot \begin{bmatrix} \underline{a} \\ 1 \end{bmatrix} + 1 \pmod 3. \tag{71}$$

There are two solutions for (70) and (71):

$$T_3(1) := [\underline{v}_3, \underline{v}_1, \underline{v}_2, \underline{w}_2, \underline{w}_3, \underline{w}_1] \cdot [\underline{v}_1, \underline{v}_2, \underline{v}_3, \underline{w}_1, \underline{w}_2, \underline{w}_3]^{-1}$$

and

$$T_3(2) := [\underline{v}_3, \underline{v}_1, \underline{v}_2, \underline{w}_3, \underline{w}_1, \underline{w}_2] \cdot [\underline{v}_1, \underline{v}_2, \underline{v}_3, \underline{w}_1, \underline{w}_2, \underline{w}_3]^{-1}.$$

In other words, $T_3(1)$ maps $\underline{v}_3 \longrightarrow \underline{v}_2 \longrightarrow \underline{v}_1 \longrightarrow \underline{v}_3$ and at the same time $\underline{w}_3 \longrightarrow \underline{w}_2 \longrightarrow \underline{w}_1 \longrightarrow \underline{w}_3$, while $T_3(2)$ has the same action on $P_3(2)$ and the reverse action $\underline{w}_3 \longleftarrow \underline{w}_2 \longleftarrow \underline{w}_1 \longleftarrow \underline{w}_3$ on $P_3(5)$. Consequently, $T_3(1)T_3(2) = T_3(2)T_3(1)$.

TABLE 2-7. THE MATRICES T_3.

$3 \cdot T_3(1)$					$3 \cdot T_3(2)$				
-2	1	-4	2	2	-2	1	2	-4	2
-1	-1	-2	4	1	-1	-1	4	-2	1
2	-1	-2	1	1	1	-2	-1	-1	2
1	-2	-1	-1	2	2	-1	1	-2	1
0	0	0	0	3	0	0	0	0	3

Step 3. Verify that $T_3(1)^3 = T_3(2)^3 = I$. This is done easily.

Step 4. They preserve the size of a 3-core solution pair: if

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ 1 \end{bmatrix} := T_3 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ 1 \end{bmatrix},$$

then $3(b_1^2 + b_2^2 - b_1 b_2) + 6(b_3^2 + b_4^2 - b_3 b_4) - (b_1 + b_2) - 2(b_3 + b_4) =$ $3(a_1^2 + a_2^2 - a_1 a_2) + 6(a_3^2 + a_4^2 - a_3 a_4) - (a_1 + a_2) - 2(a_3 + a_4)$.

Step 5. Check that they map $\mathbb{Z}^d$ into itself. For this, write $t \cdot T_3(\left[\frac{a}{1}\right]) = X\underline{a} + \underline{b}$ again. Notice that

$$X(1) \equiv -\begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} [1, 1, 2, 2] = \underline{\xi} \cdot \underline{e} \pmod 3$$

and

$$X(2) \equiv -\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} [1,1,2,2] = \underline{\xi} \cdot \underline{e} \pmod{3}.$$

In addition, $1 \cdot \underline{\xi} + \underline{b} \equiv 0 \pmod{3}$ for both matrices. Therefore

$$t \cdot T_3(\begin{bmatrix} \underline{a} \\ 1 \end{bmatrix}) = X\underline{a} + \underline{b} = \underline{\xi}\,(\underline{e} \cdot \underline{a}) \equiv \underline{\xi} \cdot 1 + \underline{b} \equiv \underline{0} \pmod{3}.$$

Step 6. There exist integers $\alpha = 1$ and $\beta = 0$ such that

$$[\underline{f}, 0]T - [\underline{f}, 0] \equiv \alpha \cdot [\underline{e}, 0] + \beta \cdot [\underline{0}, 1] = [1,1,2,2,0].$$

This leads to

$$\underline{f} \cdot T(\underline{a}) - \underline{f} \cdot \underline{a} \equiv \alpha \cdot (\underline{ea}) + \beta \equiv 1 \cdot 1 + 0 = 1 \pmod{3}.$$

Step 7. The relationship $\det T_3(1) = \det T_3(2) = 1$ shows that $T_3(\ell)$, $\ell = 1, 2$ established bijections between
$S^*\,(3(n+1)-1, f(\ell) \equiv k \pmod{3})$ and $S^*\,(3(n+1)-1, f(\ell) \equiv k+1 \pmod{3})$.

With this, we finished the proof of Theorem 5. ∎

The two cranks on $S^*(3n+2)$, found in this section can be extended to cranks on $S(3n+2)$ by the bijection of Lemma 3. Thus we have three different (see the beginning of this section) cranks on $S(3n+2)$.

§2.10. Further Investigations; a Proof of $s^*(9n+8) \equiv 0 \pmod 9$

As in the past, let

$S^*(n) := \{[\pi(1), \pi(2)] \in P_3^* \times P_3^* : |\pi(1)| + 2|\pi(2)| = n\}$

Having two cranks, we can classify the solutions using both of them.
Let $\underline{k}(3) := [k_1, k_2, k_3] = [2, 0, 1]$ and
$\underline{k}(9) := [k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9] = [8, 2, 5, 6, 0, 3, 7, 1, 4]$.
We set up a $t \times t$ matrix M_t for $t = 3, 9$. Entry (i, j) of $M_t(n)$ will count a solution $[\pi_1, \pi_2] \in S^*(n)$ if the image $\underline{a}$ of the solution in $\mathbb{Z}^4$ satisfies both $\underline{f}(1) \cdot \underline{a} \equiv k_i \pmod t$ and $\underline{f}(2) \cdot \underline{a} \equiv k_j \pmod t$.

For example, the two cranks classify the 42, $3 \cdot 42$, $3^2 \cdot 42$ solutions of $|\pi(1)| + 2|\pi(2)| = 40$, 122, 368 into the following matrices:

TABLE 2-8. EXAMPLES OF THE DUAL CLASSIFICATION.

$M_3(40)$

	2	0	1
2	1	5	6
0	5	8	5
1	6	5	1

$M_9(40)$

	8	2	5	6	0	3	7	1	4
8	0	0	0	0	0	2	0	2	2
2	0	1	0	0	2	0	0	0	0
5	0	0	0	0	0	1	0	2	0
6	0	0	0	2	0	1	0	2	1
0	0	2	0	0	2	0	2	0	0
3	2	0	1	1	0	2	0	0	0
7	0	0	0	0	2	0	1	0	0
1	2	0	2	2	0	0	0	0	0
4	2	0	0	1	0	0	0	0	0

$M_3(122)$

	2	0	1
2	12	12	18
0	12	18	12
1	18	12	12

$M_9(122)$

	8	2	5	6	0	3	7	1	4
8	0	0	1	4	0	2	2	0	3
2	0	4	2	0	0	1	0	2	3
5	1	2	2	2	1	2	3	3	2
6	4	0	2	2	3	0	1	2	2
0	0	0	1	3	2	3	0	0	1
3	2	1	2	0	3	2	0	4	2
7	2	0	3	1	0	0	4	0	2
1	0	2	3	2	0	4	0	0	1
4	3	3	2	2	1	2	2	1	2

$M_3(368)$

	2	0	1
2	42	42	42
0	42	42	42
1	42	42	42

$M_9(368)$

	8	2	5	6	0	3	7	1	4
8	6	7	7	5	5	1	3	4	4
2	7	2	2	5	1	5	4	3	4
5	7	2	2	8	6	6	4	4	12
6	5	5	8	1	5	6	5	1	6
0	5	1	6	5	8	5	1	5	6
3	1	5	6	6	5	1	5	5	8
7	3	4	4	5	1	5	2	7	2
1	4	3	4	1	5	5	7	6	7
4	4	4	12	6	6	8	2	7	2

From the symmetries of the quadratic form of Theorem 4 we can see that the general form of the matrices is

$M_3(n)$		
c	b	d
b	a	b
d	b	c

$M_9(n)$								
N	R	Q	H	E	K	Y	T	X
R	P	S	J	G	M	V	Y	Z
Q	S	O	I	F	L	Z	X	U
H	J	I	C	B	D	M	K	L
E	G	F	B	A	B	G	E	F
K	M	L	D	B	C	J	H	I
Y	V	Z	M	G	J	P	R	S
T	Y	X	K	E	H	R	N	Q
X	Z	U	L	F	I	S	Q	O

These matrices have plenty of interesting properties, let us prove some of them.

THEOREM 6.

(a) All nine entries of $M_3(9n+8)$ are equal.

(b) There is a kind of self-similarity: the central ninth of the blow-up $M_9(9n+8)$ is identical to $M_3(n)$.

(c) $M_3(3n+2)$ can appear in one of the following three forms, according to $n \pmod 3$:

$n \equiv 0$		
a'	b'	b'
b'	a'	b'
b'	b'	a'

$n \equiv 1$		
b'	b'	a'
b'	a'	b'
a'	b'	b'

$n \equiv 2$		
b'	b'	b'
b'	b'	b'
b'	b'	b'

where $a' = a + 2b$ and $b' = b + c + d$.

Let us remark that statement (a) proves (11) of §1 for $\alpha = 2$.

THEOREM 7.

(a) $M_3(2n+1) =$

d	b	c
b	a	b
c	b	d

(b) $M_9(n) = M_9(2^6(n+1) - 1)$.

To *prove* these phenomena, we shall first find injections

$U_2 : S^*(n) \longrightarrow S^*(2n+1)$ and $U_3 : S^*(n) \longrightarrow S^*(3n+2)$. This involves some search by trial and error. The winning definitions are the following.

Let $U_3(1)$ be the affine transformation that takes

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \in [P_3(2), P_3(2), P_3(4), P_3(4), P_3(0)] \tag{72}$$

into

$$\begin{bmatrix} 0 & -1 & 2 & 0 & 0 \\ -1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix} \in [P_3(8), P_3(8), P_3(14), P_3(14), P_3(2)]$$

The image was chosen so that $\underline{f}(1) \cdot U_3(1) \equiv \underline{0} \pmod 3$.

$U_3(2)$ is defined to take the same set of vectors into

$$\begin{bmatrix} 0 & -1 & 2 & 0 & 0 \\ -1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

This image was chosen so that $\underline{f}(2) \cdot U_3(2) \equiv \underline{0} \pmod 3$.

Finally, U_2 is defined to take the basis (72) into

$$\begin{bmatrix} 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \in [P_3(5), P_3(5), P_3(9), P_3(9), P_3(1)]$$

One can check (for example by Mathematica) that the application to $\underline{a}$ of the resulting matrices

$U_3(1) =$				
0	-1	2	0	0
-1	0	0	2	0
-1	0	0	-1	1
0	-1	-1	0	1
0	0	0	0	1

$U_3(2) =$				
0	-1	2	0	0
-1	0	0	2	0
0	-1	-1	0	1
-1	0	0	-1	1
0	0	0	0	1

$U_2 =$				
0	0	0	-2	1
0	0	-2	0	1
0	1	0	0	0
1	0	0	0	0
0	0	0	0	1

do increase the value of the quadratic expression $3(a_1^2 + a_2^2 - a_1a_2) + 6(a_3^2 + a_4^2 - a_3a_4) - (a_1 + a_2) - 2(a_3 + a_4)$ from n to $3n + 2$ or $2n + 1$ for every n.

Step 2. Let us compute $\underline{f}(1) \cdot U^x$ and $\underline{f}(2) \cdot U^x$. Since $\det U_2 = 2^2, \det U_3(1) = 3^2, \det U_3(2) = -3^2$, the inverse matrices ($x = -1$) do exist.

With the abbreviations $\underline{v}_1 := -\underline{f}(2) + \underline{v}_4$, $\underline{v}_2 := \underline{f}(1) + \underline{v}_4$, $\underline{v}_3 := \underline{f}(1) + \underline{v}_5$, where $\underline{v}_4 := [0, 0, -1, 1]$ and $\underline{v}_5 := [-1, 1, 0, 0]$, we have the following tables (the fifth coordinates, if zero, are omitted):

TABLE 2-9. THE EFFECTS OF ITERATING $U_3(1)$.

$\cdot$ (mod 9)	$U_3(1)^{-1}$	$U_3(1)^0$	$U_3(1)^1$	$U_3(1)^2$	$U_3(1)^3$	$U_3(1)^4$	$U_3(1)^5$	$U_3(1)^6$	$U_3(1)^7$	$U_3(1)^8$	$U_3(1)^9$
$\underline{f}(1)$	$-\frac{1}{3}\underline{f}(1) + \underline{v}_5$	$\underline{f}(1)$	$3\underline{v}_4$	$-3\underline{f}(2)$	$-3\underline{v}_3$	$-3\underline{f}(2)$	$6\underline{v}_3$	$3\underline{v}_3$	$3\underline{f}(2)$	$-6\underline{v}_3$	$-3\underline{v}_3$
$\underline{f}(2)$	$-\underline{v}_4$	$\underline{f}(2)$	$\underline{v}_3$	$\underline{f}(2) - 3\underline{v}_4$	$-2\underline{v}_3$	$-\underline{v}_3$	$-\underline{f}(2) + 3\underline{v}_4$	$2\underline{v}_3$	$\underline{v}_3$	$\underline{f}(2) - 3\underline{v}_4$	

TABLE 2-10. THE EFFECTS OF ITERATING $U_3(2)$.

$\cdot$	$U_3(2)^{-1}$	$U_3(2)^0$	$U_3(2)^1$	$U_3(2)^2$	$U_3(2)^3$	$U_3(2)^4$	$U_3(2)^5$
$\underline{f}(1)$	$-\frac{1}{3}\underline{f}(2) + \underline{v}_5$	$\underline{f}(1)$	$\underline{v}_3$	$3\underline{f}(1)$	$3\underline{v}_3$	$9\underline{f}(1)$	$9\underline{v}_3$
$\underline{f}(2)$	$\underline{v}_4$	$\underline{f}(2)$	$3\underline{v}_4$	$3\underline{f}(2)$	$9\underline{v}_4$	$9\underline{f}(2)$	$27\underline{v}_4$

TABLE 2-11. THE EFFECTS OF ITERATING U_2.

$\cdot$	U_2^{-1}	U_2^0	U_2^1	U_2^2	U_2^3	U_2^4	U_2^5	U_2^6	U_2^7
$\underline{f}(1)$	$-\frac{1}{2}\underline{v}_1$	$\underline{f}(1)$	$\underline{v}_1$	$-2\underline{f}(1)$	$-2\underline{v}_1$	$4\underline{f}(1)$	$4\underline{v}_1$	$-8\underline{f}(1)$	$-8\underline{v}_1$
$\underline{f}(2)$	$-\frac{1}{2}\underline{v}_2$	$\underline{f}(2)$	$\underline{v}_2$	$-2\underline{f}(2)$	$-2\underline{v}_2$	$4\underline{f}(2)$	$4\underline{v}_2$	$-8\underline{f}(2)$	$-8\underline{v}_2$

TABLE 2-12. $T_3(2)^x U_3(1)$.

$\cdot$ (mod 3)	$T_3(2)^0U_3(1)$	$T_3(2)^1U_3(1)$	$T_3(2)^2U_3(1)$
$\underline{f}(1)$	0	$[-\underline{e},1]$	$[\underline{e},-1]$
$\underline{f}(2)$	$-\underline{f}(2)$	$[-\underline{f}(2),1]$	$[-\underline{f}(2),-1]$

TABLE 2-13. $T_3()^x U_3(1)^2$.

$\cdot$ (mod 3)	$T_3(2)^2U_3(1)^2$	$T_3(2)^1U_3(1)^2$	$U_3(1)^2$	$T_3(1)^1U_3(1)^2$	$T_3(1)^2U_3(1)^2$
$\underline{f}(1)$	0	0	0	1	2
$\underline{f}(2)$	$[\underline{f}(2),2]$	$[\underline{f}(2),1]$	$\underline{f}(2)$	$\underline{f}(2)$	$\underline{f}(2)$

An interesting fact from Table 9, although we do not need here is that $\underline{f}(1)U_3(1)^{x+2} \equiv -3\underline{f}(2)U_3(1)^x \pmod 9$.

PROOF OF THEOREM 7:

To prove (a) of Theorem, note that $\underline{v}_1 \equiv -\underline{f}(1) \pmod 3$ and $\underline{v}_2 \equiv \underline{f}(2) \pmod 3$.

Therefore if $M_3(n) = \begin{bmatrix} c_1 & b_1 & d_1 \\ b_2 & a & b_3 \\ d_2 & b_4 & c_2 \end{bmatrix}$ then $M_3(2n+1) = \begin{bmatrix} d_2 & b_4 & c_2 \\ b_2 & a & b_3 \\ c_1 & b_1 & d_1 \end{bmatrix}$.

The 6-cycle in the apparent orbit of U_2 in M_9 (Table 11) proves (b). ∎

PROOF OF THEOREM 6:

The identity

$$\frac{[\underline{f}(1)U_3(2)^2, \underline{f}(2)U_3(2)^2]}{9} = \frac{[\underline{f}(1), \underline{f}(2)]}{3},$$

(which can be read from Table 10) proves (b).

Note that $\underline{v}_3 \equiv -\underline{f}(2) \pmod 3$. Table 9 tells us that a solution in $[\underline{f}(1), \underline{f}(2)]$

of M_3 is mapped into $[0, -\underline{f}(2)] \in M_3(3n+2)$ by $U_3(1)$.

Therefore, if $M_3(n) = \begin{bmatrix} c_1 & b_1 & d_1 \\ b_2 & a & b_3 \\ d_2 & b_4 & c_2 \end{bmatrix}$, then the image of it by $U_3(1)$ is

$$U_3(1)M_3(3n) = \begin{bmatrix} 0 & 0 & 0 \\ c_2+b_3+d_1 & b_4+a+b_1 & d_2+b_2+c_1 \\ 0 & 0 & 0 \end{bmatrix}. \tag{73}$$

Regard Table 12. Since $\underline{e} \cdot \underline{a} \equiv n \pmod 3$, one iteration by matrix $T_3(2)$ moves a solution in (73) by $(n+1) \pmod 3$ vertically and by 1 to the right:

$n \equiv 0$			
a'	b_2'	b_1'	$T_3(2)^2U_3(1)$
b_1'	a'	b_2'	$T_3(2)^0U_3(1)$
b_2'	b_1'	a'	$T_3(2)^1U_3(1)$

$n \equiv 1$			
b_2'	b_1'	a'	$T_3(2)^1U_3(1)$
b_1'	a'	b_2'	$T_3(2)^0U_3(1)$
a'	b_2'	b_1'	$T_3(2)^2U_3(1)$

$n \equiv 2$		
0	0	0
b'=a'	a'	b'
0	0	0

This proved statement (c) for $n \equiv 0, 1$, but showed only $a' = b'$ for $n \equiv 2$. Of course, we know from $3(3n+2)+2 = 9n+8$ that the case $n \equiv 2$ of (c) and (a) are the same statements.

Consider Table 13. Let the bullet, $\bullet$ mark the position of a possible solution in M_3, if the exponent x is zero. The arrows indicate the direction of the change as x increases. The next two diagrams in M_3 give a proof of (a):

$$T_3(1)^x U_3(1)^2 : \begin{array}{ccc} \circ & \circ & \circ \\ \downarrow & \downarrow & \downarrow \\ \bullet & \bullet & \bullet \\ \downarrow & \downarrow & \downarrow \\ \circ & \circ & \circ \end{array} \qquad T_3(2)^x U_3(1)^2 : \begin{array}{ccccc} \times & & \times & & \times \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \times & & \times & & \times \end{array}$$

Since $\det T_3 = 1$, the maps above are all one-to-one, therefore the diagrams above prove (a). ∎

CHAPTER 3

A CLASSIFICATION OF UNRESTRICTED PARTITIONS

To give a combinatorial explanation to Ramanujan's congruences

$$p(5n+4) \equiv 0 \pmod 5 \text{ and } p(7n+5) \equiv 0 \pmod 7 ,$$

Dyson introduced the *rank*, which is (the length of the largest part) -(number of parts). In contrast to the rank, we introduce the notion of *frame* to denote the union of the first row and first column in the Ferrers graph of the partition. The *frame size* is the (length of the largest part) + (number of parts) -1. If we are more relaxed with our usage, we use the word frame for frame size. Note that the frame and the rank are of opposite parity. Let $p_r(n)$ be the number of partitions of n with frame size r. Obviously, $\sum_r p_r(n) = p(n)$. We will examine a few basic characteristics of this new classification. By peeling the frame off, we get to a partition of the integer $n - r$. By reversing this process, we can see that the sequence of successive frame sizes and ranks determine the partition.

LEMMA 1. *The row sums are consecutive powers of two,*

$$\sum_n p_r(n) = 2^{r-1}.$$

PROOF: Suppose we have a partition of n with frame size r. This means that if the largest part is of size j, where $1 \leq j \leq r$, then the number of parts is $r - j + 1$. This partition can have $j - 1$ horizontal and $r - j$ vertical strokes on its boundary connecting the legs of the Γ of its Ferrers graph. This gives $\binom{r-1}{j-1}$ choices, and $\sum_{j=1}^{r} \binom{r-1}{j-1} = 2^{r-1}$. ∎

TABLE 3-1. $p_r(n)$, WHERE (n, r)= (PARTITIONED INTEGER, FRAME SIZE)

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22
22																						22
21																					21	19
20																				20	18	34
19																			19	17	32	45
18																		18	16	30	42	66
17																	17	15	28	39	61	79
16																16	14	26	36	56	72	106
15															15	13	24	33	51	65	95	117
14														14	12	22	30	46	58	84	102	136
13													13	11	20	27	41	51	73	87	114	134
12												12	10	18	24	36	44	62	72	92	104	120
11											11	9	16	21	31	37	51	57	70	74	78	77
10										10	8	14	18	26	30	40	42	48	44	46	40	40
9									9	7	12	15	21	23	29	27	26	23	21	15	13	7
8								8	6	10	12	16	16	18	12	12	8	6	2	2		
7							7	5	8	9	11	9	7	4	3	1						
6						6	4	6	6	6	2	2										
5					5	3	4	3	1													
4				4	2	2																
3			3	1																		
2		2																				
1	1																					
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22

Lemma 2. *$p_r(n) > 0$ iff r is between roughly $\sqrt{n}$ and n.*

In detail, $\{(n,r) : p_r(n) > 0\} = \{C_\ell\}_{\ell=0}^{\infty}$, where C_ℓ are half lines with slope 45°, which we call cuts. To avoid writing 0 indices everywhere, let us denote the endpoints of these cuts by (n,r) too, i.e.

$$C_\ell = \{(n+d, r+d) : \; n-r=\ell, \; d \geq 0\}.$$

Place n between two consecutive squares, k^2 and $(k+1)^2$.

(a) If $n = k^2 + m$, where $1 \leq m \leq k$, then $r = 2k$.

(b) If $n = k^2 + k + m$, where $1 \leq m \leq k+1$, then $r = 2k+1$. ∎

Lemma 3. *Let ϵ be 0 or 1.*

If $(n,r) = (k^2 + \epsilon \cdot k + m, 2k+\epsilon)$, then

$$(n+r+2, r+2) = ((k+1)^2 + \epsilon \cdot (k+1) + (m+1), \; 2(k+1)+\epsilon)$$

$$(n-r, r-2) = ((k-1)^2 + \epsilon \cdot (k-1) + (m-1), \; 2(k-1)+\epsilon)$$

$$(n-2r+2, r-4) = ((k-2)^2 + \epsilon \cdot (k-2) + (m-2), \; 2(k-2)+\epsilon) \; ∎$$

Fix (n,r) as the endpoint of a cut , i.e. $p_r(n) > 0$, but $p_{r-1}(n-1) = 0$. We gave the relationship between r and n in Lemma 3., and now we are going to ascend along the cut.

We are interested in the numbers $\{p_{r+d}(n+d) : d \geq 0\}$. Let us introduce the shorthand notation $p_k^* := p_k(n-r)$ for the partitions of $(n-r)$. We have the following identities:

$$p_r(n) = 1 \cdot p_{r-2}^*,$$

$$p_{r+1}(n+1) = 2 \cdot p_{r-2}^* + 1 \cdot p_{r-1}^*,$$

$$p_{r+d}(n+d) = (d+1)\cdot p^*_{r-2} + d\cdot p^*_{r-1} + \ldots + 1\cdot p^*_{r-2+d},$$

$$\textit{where } 0 \le d \le n-2r+2,$$

$$p_{n-r+2}(2n-2r+2) = (n-2r+3)\cdot p^*_{r-2} + (n-2r+2)\cdot p^*_{r-1} + \ldots + 2\cdot p^*_{n-r-1} + 1\cdot p^*_{n-r},$$

$$p_{r+d}(n+d) = (d+1)\cdot p^*_{r-2} + d\cdot p^*_{r-1} + \ldots + (d-n+2r)\cdot p^*_{n-r-1} + (d-n+2r-1)\cdot p^*_{n-r},$$

$$\textit{where } d \ge n-2r+1.$$

We can summarize the previous identities in one lemma.

LEMMA 4. *Let $(x)_+$ be max$\{x,0\}$. For any $d \ge 0$ we have the identity*

$$p_{r+d}(n+d) = \sum_{k=0}^{n-2r+2} (d+1-k)_+ \cdot p^*_{r-2+k}.$$

∎

COROLLARY. $p_{r+1}(n+1) - p_r(n) = p(r-n)$ *iff* $r \ge \frac{n+1}{2}$

PROOF: Notice that the differences $p_{r+d+1}(n+d+1) - p_{r+d}(n+d)$ are all equal to $p^*_{r-2} + p^*_{r-1} + \ldots + p^*_{n-r-1} + p^*_{n-r} = p(n-r)$, as soon as $d \ge n-2r+1$. ∎

LEMMA 5. *If r is even, then $p_r(n)$ is even as well.*

PROOF: We want to prove that the lemma is true along every cut C_ℓ. The proof uses induction on ℓ. Suppose the lemma is true for $\ell = 1, 2, \ldots, n-r-1$. If $r > 2$, we have $(n-r)-(r-2) < n-r$, therefore the lemma is true at each entry in the column with horizontal coordinate $n-r$. If r is even, then $p_r(n) = p_{r-2}(n-r)$ is even, $p_{r+2}(n+2) = p_r(n) + p_r(n-r)$, $p_{r+4}(n+4) = p_{r+2}(n+2) + p_{r+2}(n-r)$, etc. are all even numbers. If r is odd, then $p_{r+1}(n+1) = 2p_{r-2}(n-r) + p_{r-1}(n-r)$ is even and

now $p_{r+3}(n+3) = p_{r+1}(n+1) + p_{r+1}(n-r)$, $p_{r+5}(n+5) = p_{r+3}(n+3) + p_{r+3}(n-r)$ are all even. Finally, If $r \geq \frac{n+1}{2}$, we have $p_{r+2}(n+2) = p_r(n) + 2p(n-r)$, therefore the parity cannot change. ∎

CHAPTER 4

AN EXTENSION OF BAILEY'S LEMMA

§4.1. The Classical Case

This introduction is intended to be short. The reader is encouraged to consult [Lil93] about applications and for further references.

Let the sequences $\{a_r\}_0^\infty$ and $\{a_n\}_0^\infty$ be connected by the identity

$$A_n = \sum_{r=0}^{n} \frac{a_r}{(q)_{n-r}(cq)_{n+r}}. \tag{a1}$$

These sequences are called *Bailey pairs.* (The equations of the first two sections are numbered (a1), (a2), etc., while the formulas quoted from §3 [the reference section] are numbered (1), (2), etc.).

Let us define two new sequences.

$$a'_r := \frac{(\gamma_1, \gamma_2)_r}{\left(\frac{cq}{\gamma_1}, \frac{cq}{\gamma_2}\right)_r} \cdot \left(\frac{cq}{\gamma_1\gamma_2}\right)^r a_r \tag{a2}$$

and

$$A'_N := \frac{1}{\left(\frac{cq}{\gamma_1}, \frac{cq}{\gamma_2}\right)_N} \sum_{n=0}^{N} \frac{\left(\frac{cq}{\gamma_1\gamma_2}\right)_{N-n}}{(q)_{N-n}} \cdot (\gamma_1, \gamma_2)_n \cdot \left(\frac{cq}{\gamma_1\gamma_2}\right)^n A_n. \tag{a3}$$

Then $\underline{A}'$ and $\underline{a}'$ have the same relationship as (a1), namely

$$A'_n = \sum_{r=0}^{n} \frac{a'_r}{(q)_{n-r}(cq)_{n+r}}.$$

In other words, we have the following lemma.

LEMMA 1 (BAILEY LEMMA).

$$\sum_{n=0}^{N} \frac{\left(\frac{cq}{\gamma_1\gamma_2}\right)_{N-n}}{(q)_{N-n}} \cdot (\gamma_1,\gamma_2)_n \cdot \left(\frac{cq}{\gamma_1\gamma_2}\right)^n A_n =$$

$$= \left(\frac{cq}{\gamma_1},\frac{cq}{\gamma_2}\right)_N \cdot \sum_{n=0}^{N} \frac{(\gamma_1,\gamma_2)_n}{\left(\frac{cq}{\gamma_1},\frac{cq}{\gamma_2}\right)_n} \cdot \frac{\left(\frac{cq}{\gamma_1\gamma_2}\right)^n}{(q)_{N-n}(cq)_{N+n}} a_n.$$

COROLLARIES.

$$\sum_{n=0}^{\infty} (\gamma_1,\gamma_2)_n \left(\frac{cq}{\gamma_1\gamma_2}\right)^n A_n = \frac{\left(\frac{cq}{\gamma_1},\frac{cq}{\gamma_2}\right)_\infty}{\left(cq,\frac{cq}{\gamma_1\gamma_2}\right)_\infty} \cdot \sum_{n=0}^{\infty} \frac{(\gamma_1,\gamma_2)_n}{\left(\frac{cq}{\gamma_1},\frac{cq}{\gamma_2}\right)_n} \cdot \left(\frac{cq}{\gamma_1\gamma_2}\right)^n a_n. \tag{a4}$$

$$\sum_{n=0}^{N} \frac{c^n q^{n^2}}{(q)_{N-n}} A_n = \sum_{n=0}^{N} \frac{c^n q^{n^2}}{(q)_{N-n}(cq)_{N+n}} a_n \tag{a5}$$

$$\sum_{n=0}^{\infty} c^n q^{n^2} A_n = \frac{1}{(cq)_\infty} \sum_{n=0}^{\infty} c^n q^{n^2} a_n \tag{a6}$$

The following diagram shows how the corollaries were obtained from the lemma:

$$\begin{array}{ccc} (Lemma 1) & \xrightarrow{\gamma 1,\gamma 2\to\infty} & (a5) \\ \big\downarrow{\scriptstyle N\to\infty} & & \big\downarrow{\scriptstyle N\to\infty} \\ (a4) & \xrightarrow{\gamma 1,\gamma 2\to\infty} & (a6) \end{array}$$

G. Andrews made several observations. First of all, we can iterate the process of creating new pairs by (a1) and (a2). Second, we can iterate it backward too. This way we obtain the following chain:

$$\begin{tikzcd}
{} & {} \\
\underline{a}^{(-2)} \arrow[u] \arrow[r] & \underline{A}^{(-2)} \arrow[u] \\
\underline{a}^{(-1)} \arrow[u] \arrow[r] & \underline{A}^{(-1)} \arrow[u] \\
\underline{a} \arrow[u] \arrow[r] \arrow[d] & \underline{A} \arrow[u] \arrow[d] \\
\underline{a}^{(1)} \arrow[r] \arrow[d] & \underline{A}^{(1)} \arrow[d] \\
\underline{a}^{(2)} \arrow[r] \arrow[d] & \underline{A}^{(2)} \arrow[d] \\
{} & {}
\end{tikzcd}$$

The iterated Bailey lemma has led to many applications in additive number theory, combinatorics, special functions and mathematical physics. For example, the Rogers-Ramanujan identities can be proved by applying two iterations (see the article [Lil93] for further references).

The Bailey lemma can be cast in a matrix form. Define the matrices P_c, Q and matrix ∂Q consisting of diagonal elements of Q:

$$P_c\,(n,r) := \frac{1}{(q)_{n-r}\,(cq)_{n+r}}, \tag{a7}$$

$$Q\,(n,r) := \frac{(\gamma_1,\gamma_2)_r}{\left(\frac{cq}{\gamma_1},\frac{cq}{\gamma_2}\right)_n} \cdot \frac{\left(\frac{cq}{\gamma_1\gamma_2}\right)_{n-r}}{(q)_{n-r}} \cdot \left(\frac{cq}{\gamma_1\gamma_2}\right)^r \tag{a8}$$

and

$$\partial Q = \frac{(\gamma_1,\gamma_2)_r}{\left(\frac{cq}{\gamma_1},\frac{cq}{\gamma_2}\right)_r} \cdot \left(\frac{cq}{\gamma_1\gamma_2}\right)^r \cdot \delta_{nr}. \tag{a9}$$

Bailey's lemma expresses the relationship

$P_c = Q \cdot P_c \cdot (\partial Q)^{-1}$ or $\partial Q = (P_c)^{-1} \cdot Q \cdot P_c.$

$$\begin{tikzcd} \underline{a} \arrow[r, "P_c"] \arrow[d, "\partial Q"] & \underline{A} \arrow[d, "Q"] \\ \underline{a'} \arrow[r, "P_c"] & \underline{A'} \end{tikzcd}$$

The inverse of matrix P_c is

$$P_c^{-1}(n,r) = \frac{1-cq^{2n}}{1-c} \cdot \frac{(c)_{n+r}}{(q)_{n-r}} \cdot (-1)^{n-r}\, q^{\binom{n-r}{2}}. \tag{a10}$$

D. Bressoud [Bre83] has given an elegant proof of (a10) by introducing the matrices $L_{b,c}$ and showing that

$$L_{b,c} \cdot L_{c,d} = L_{c,d},$$

where

$$L_{b,c}(n,r) = \frac{\left(\frac{b}{c}\right)_{n-r} (b)_{n+r}}{(q)_{n-r}(cq)_{n+r}} \cdot \frac{1-cq^{2r}}{1-c} c^{n-r}.$$

In particular, $L_{b,c} \cdot L_{c,b} = I$. Notice that $P_c = D_1 \cdot L_{0,c} \cdot D_2$, where D_1 and D_2 are diagonal matrices. Therefore

$$P_c^{-1} = D_2^{-1} \cdot L_{c,0} \cdot D_1^{-1} = D_2^{-1} \cdot \lim_{\epsilon \to 0} L_{c,\epsilon} \cdot D_1^{-1},$$

which is equal to (a10) using (1) from §3.

The question arises what happens if we change the parameter c of matrix P_c at iterated applications of Bailey's lemma. If we are given parameters c, c' of matrix P_c (defined in (a7)) and matrix Q (defined in (a8)) what matrix Q' would make the following diagram commutative?

$$\begin{tikzcd} \underline{a} \arrow[r, "P_c"] \arrow[d, "Q'"] & \underline{A} \arrow[d, "Q"] \\ \underline{a'} \arrow[r, "P_{c'}"] & \underline{A'} \end{tikzcd}$$

Agarwal et al. [Aga87] answered the question for $c' = \frac{c}{q}$. In the next section, we answer the question for any c', by analyzing their proof. We will also be able to get a result with several extra parameters added to the definition of matrix Q. Moreover, Q can be a finite product of such matrices.

§4.2. The Extension

THEOREM 1 (EXTENDED BAILEY'S LEMMA).

$$\begin{tikzcd}
\underline{a} \arrow[r, "P_c"] \arrow[d] & \underline{A} \arrow[d, "Q_k"] \\
{} \arrow[d] & {} \arrow[d, "\vdots"] \\
{} \arrow[d, "Q'"] & {} \arrow[d, "Q_1"] \\
\underline{a'} \arrow[r, "P_{c'}"] & \underline{A'}
\end{tikzcd}$$

Let the $(n,r)^{th}$ element of the following matrices be defined by

$$P_c\,(n,r) := \frac{1}{(q)_{n-r}\,(cq)_{n+r}}$$

$$Q_i\,(n,r) := \frac{(\gamma_{1i}, \gamma_{2i})_r}{(\gamma_{3i}, \gamma_{4i})_n} \frac{(\gamma_{5i}q)_{n-r}}{(q)_{n-r}} \left(\gamma_{6i}q\right)^r .$$

Here $i = 1, 2, \ldots, k$, where $k \geq 1$. These matrices are infinite to the right and down, since $n, r \geq 0$. They are also lower triangular, by property (3b) of §3 *Let's form now matrices from the diagonals of Q_i's:*

$$\partial Q_i\,(n,r) := \frac{(\gamma_{1i}, \gamma_{2i})_n}{(\gamma_{3i}, \gamma_{4i})_n} \left(\gamma_{6i}q\right)^n \delta_{nr}$$

Now, if $\gamma_{5i} = \gamma_{6i}$ *and if*

$$\frac{\gamma_{3i}\gamma_{4i}}{\gamma_{5i}q} = \gamma_{1i}\gamma_{3i} = \gamma_{2i}\gamma_{4i} = c'q \tag{a11}$$

then

$$Q'(n,r) := \left(P_{c'}^{-1}Q_1 \ldots Q_k P_c\right)(n,r) =$$
$$\prod_{i=1}^{k} \partial Q_i(n,n) \cdot \frac{1-c'q^{2n}}{1-c'}\,(c')_{n+r}\left(\frac{c'}{c}\right)_{n-r}\left(cq^{n+r}\right)^{n-r} \cdot P_c(n,r),$$

where $Q'\underline{a} = \underline{a}'$.

The duo $\underline{a}$ and $\underline{A}$ on the diagram is called a *Bailey pair,* In applications, we apply matrix $Q_1 \cdot \ldots \cdot Q_k$ to the vector $\underline{A}$ on one hand, and matrix $P_{c'}Q'$ to the vector $\underline{a}$ on the other hand and take the $n_0{}^{th}$ coordinates on both sides to get the identity

$$\sum_{n_k=0}^{n_0}\sum_{n_{k-1}=n_k}^{n_0}\cdots\sum_{n_1=n_2}^{n_0} Q_1(n_0,n_1)\ldots Q_k(n_{k-1},n_k)\,\underline{A}(n_k) =$$
$$\sum_{n_2=0}^{n_0}\sum_{n_1=n_2}^{n_0} P_{c'}(n_0,n_1)\,Q'(n_1,n_2)\,\underline{a}(n_2).$$

COROLLARY.

$$\text{if } c' = c \text{ then } Q'(n,r) = \prod_{i=1}^{k} \partial Q_i(n,n) \cdot \delta_{nr}$$
$$\text{if } c' = \frac{c}{q} \text{ then } Q'(n,r) = \prod_{i=1}^{k} \partial Q_i(n,n) \cdot \frac{1-cq^{2r}}{1-c} \cdot \begin{cases} 1, & \text{if } r = n \\ -cq^{2r}, & \text{if } r = n-1. \end{cases}$$

For each i, the conditions of (a11) give us 3=7-4 or, rather 3=6-3 degrees of freedom tochoose the parameters $\gamma_1, \ldots, \gamma_6$ and c'. In the literature we have seen only the following choice of parameters:

$$\gamma_3 = \frac{c'q}{\gamma_1}, \gamma_4 = \frac{c'q}{\gamma_2}, \gamma_5 = \gamma_6 = \frac{c'}{\gamma_1\gamma_2} \tag{a12}$$

The case $(k = 1, c' = c)$ is the classical Bailey's lemma, the case $(k > 1, c' = c)$ gives Bailey's lemma with iteration and the case $(k = 1, c' = \frac{c}{q})$ can be found in the article [Aga87]. The proof in that article actually gives the theorem for any c', if $k = 1$ and if γ's are chosen by (a12).

Let's write down a list of possible choices of parameters:

TABLE 4-1. CHOICE OF PARAMETERS

γ_1	γ_2	γ_3	γ_4	$\gamma_5 = \gamma_6$	c'
-	-	$\frac{c'q}{\gamma_1}$	$\frac{c'q}{\gamma_2}$	$\frac{c'}{\gamma_1\gamma_2}$	-
$\frac{c'q}{\gamma_3}$	$\frac{c'q}{\gamma_4}$	-	-	$\frac{\gamma_3\gamma_4}{c'q^2}$	-
-	$\frac{c'q}{\gamma_4}$	$\frac{c'q}{\gamma_1}$	-	$\frac{\gamma_4}{\gamma_1 q}$	-
-	$\frac{c'}{\gamma_1\gamma_5 q}$	$\frac{c'q}{\gamma_1}$	$\gamma_1\gamma_5 q^2$	-	-
$\frac{c'q}{\gamma_3}$	$\frac{\gamma_3}{\gamma_5 q}$	-	$\frac{c'q^2\gamma_5}{\gamma_3}$	-	-
-	-	$\gamma_2\gamma_5 q$	$\gamma_1\gamma_5 q$	-	$\gamma_1\gamma_2\gamma_5$
$\frac{\gamma_4}{\gamma_5 q}$	$\frac{\gamma_3}{\gamma_5 q}$	-	-	-	$\frac{\gamma_3\gamma_4}{\gamma_5 q^2}$
-	$\frac{\gamma_3}{\gamma_5 q}$	-	$\gamma_1\gamma_5 q$	-	$\frac{\gamma_1\gamma_3}{q}$
-	-	-	$\frac{\gamma_1\gamma_3}{\gamma_2}$	$\frac{\gamma_3}{\gamma_2 q}$	$\frac{\gamma_1\gamma_3}{q}$
-	$\frac{\gamma_1\gamma_3}{\gamma_4}$	-	-	$\frac{\gamma_4}{\gamma_1 q}$	$\frac{\gamma_1\gamma_3}{q}$

LEMMA 2. *Let*

$$M_0 := n_1 + \cdots + n_\ell, \quad M_1 := n_2 + \cdots + n_\ell, \ldots \quad M_{\ell-1} := n_\ell, \qquad M_\ell := 0$$

$$N_0 := n_0 - M_0, \qquad N_1 := n_0 - M_1, \ldots \qquad N_{\ell-1} := n_0 - M_{\ell-1}, \quad N_\ell := n_0$$

If $n_1, \ldots, n_\ell, N_0 \geq 0$ *then*

$$(-1)^{N_0} q^{\binom{N_0}{2}} \cdot \frac{1}{(q)_{N_0}} = (-1)^{n_0} q^{\binom{n_0}{2}} \cdot \frac{1}{(q)_{n_0}} \cdot \left\{ \left(\frac{1}{q^{N_1}}\right)_{n_1} q^{n_1} \right\} \cdot \ldots \cdot \left\{ \left(\frac{1}{q^{N_\ell}}\right)_{n_\ell} q^{n_\ell} \right\}$$

PROOF: Use (8a) ℓ times for $N = N_1, \ldots, N_\ell$ and $r = n_1, \ldots, n_\ell$ respectively to obtain

$$\frac{1}{(q)_{N_0}} = \left\{ (-1)^{n_1} q^{N_1 n_1 - \binom{n_1}{2}} \cdot \left(\frac{1}{q^{N_1}}\right)_{n_1} \right\} \cdot \ldots \cdot \left\{ (-1)^{n_\ell} q^{N_\ell n_\ell - \binom{n_\ell}{2}} \cdot \left(\frac{1}{q^{N_\ell}}\right)_{n_\ell} \right\} \cdot \frac{1}{(q)_{n_0}}$$

then use (4) ℓ times to get

$$q^{\binom{n_0 - (n_1 + \cdots + n_\ell)}{2}} = \left\{ q^{\binom{n_1}{2} - N_1 n_1} \cdot q^{n_1} \right\} \cdot \ldots \cdot \left\{ q^{\binom{n_\ell}{2} - N_\ell n_\ell} \cdot q^{n_\ell} \right\} \cdot q^{n_0}$$

and multiply the two identities together. ∎

PROOF OF THE THEOREM: In the proof letter ℓ will denote $k+2$ of the statement. One can see that if $T_1, \ldots, T_\ell$ are lower triangular matrices, then, using the substitutions $n_{\ell-1} \leftarrow M_{\ell-2}, \ldots, n_1 \leftarrow M_0$, we get

$$(T_1 \cdot \ldots \cdot T_\ell)(n_0, n_\ell) = \sum_{n_{\ell-1}=n_\ell}^{n_0} \cdots \sum_{n_1=n_2}^{n_0} T_1(n_0, n_1) \cdot \ldots \cdot T_\ell(n_{\ell-1}, n_\ell) =$$

$$= \sum_{n_{\ell-1}=0}^{N_{\ell-1}} \cdots \sum_{n_1=0}^{N_1} T_1(n_0, M_0) T_2(M_0, M_1) \cdot \ldots \cdot T_\ell(M_{\ell-2}, M_{\ell-1}) \,.$$

We also know the inverse of matrix $P_{c'}$ by (a10). Thus we have

$$P_{c'}^{-1}(n_0, M_0) = \frac{1-c'q^{2n_0}}{1-c'} \cdot \frac{(c')_{n_0+M_0}}{(q)_{n_0-M_0}} \cdot (-1)^{n_0-M_0} q^{\binom{n_0-M_0}{2}} \overset{\text{(Lemma),also (5),(6)}}{=\!=}$$

$$= \left\{ (-1)^{n_0} q^{\binom{n_0}{2}} \frac{\left(c', q\sqrt{c'}, -q\sqrt{c'}\right)_{n_0}}{\left(q, \sqrt{c'}, -\sqrt{c'}\right)_{n_0}} \right\} \cdot$$

$$\cdot \left\{ \left(\frac{1}{q^{N_1}}, c'q^{n_0+M_1}\right)_{n_1} q^{n_1} \right\} \cdot \ldots \cdot \left\{ \left(\frac{1}{q^{N_\ell}}, c'q^{n_0+M_\ell}\right)_{n_\ell} q^{n_\ell} \right\}$$

$$=: B_0 B_1 \cdot \ldots \cdot B_\ell$$

We can break down $Q_i(M_{i-1}, M_i)$ too, where $i = 1, 2, \ldots, \ell-2$:

$$Q_i(M_{i-1}, M_i) = \frac{(\gamma_{1i}, \gamma_{2i})_{M_i}}{(\gamma_{3i}, \gamma_{4i})_{M_{i-1}}} \cdot \frac{(\gamma_{5i}q)_{n_i}}{(q)_{n_i}} (\gamma_{6i}q)^{M_i} =$$

$$= \frac{(\gamma_{5i}q)_{n_i}}{\left(q, \gamma_{3i}q^{M_i}, \gamma_{4i}q^{M_i}\right)_{n_i}} \cdot \frac{(\gamma_{1i}, \gamma_{2i})_{M_i}}{(\gamma_{3i}, \gamma_{4i})_{M_i}} (\gamma_{6i}q)^{M_i} =: G_i \cdot D_i$$

Notice that

$$P_c(n_{\ell-1} + n_\ell, n_\ell) = \frac{1}{\left(q, cq^{2n_\ell+1}\right)_{n_{\ell-1}}} \cdot \frac{1}{(cq)_{2n_\ell}}$$

Therefore

$$\left(P_{c'}^{-1} Q_1 \ldots Q_{\ell-2} P_c\right)(n_0, n_\ell) =$$

$$= \sum_{n_{\ell-1}=0}^{N_{\ell-1}} \sum_{n_{\ell-2}=0}^{N_{\ell-2}} \ldots \sum_{n_1=0}^{N_1} P_{c'}^{-1}(n_0, M_0) Q_1(M_0, M_1) \ldots Q_{\ell-2}(M_{\ell-3}, M_{\ell-2}) P_c(M_{\ell-2}, M_{\ell-1})$$

$$B_0 B_\ell \sum_{n_{\ell-1}=0}^{N_{\ell-1}} P_c B_{\ell-1} D_{\ell-2} \sum_{n_{\ell-2}=0}^{N_{\ell-2}} G_{\ell-2} B_{\ell-2} \ldots D_2 \sum_{n_2=0}^{N_2} G_2 B_2 D_1 \sum_{n_1=0}^{N_1} G_1 B_1,$$

since M_i does not depend on $\{n_0, n_1, \ldots, n_i\}$! Let us start to calculate this multiple sum from right to left:

$$\sum_{n_1=0}^{N_1} G_1 B_1 = \sum_{n_1=0}^{N_1} \frac{\left(c'q^{n_0+M_1}, \gamma_{51}q, \frac{1}{q^{N_1}}\right)_{n_1}}{\left(q, \gamma_{31}q^{M_1}, \gamma_{41}q^{M_1}\right)_{n_1}} q^{n_1} \overset{(9)}{=\!=} \frac{\left(\frac{\gamma_{31}}{c'} \frac{1}{q^{n_0}}, \frac{c'q^{M_1+1}}{\gamma_{41}}\right)_{N_1}}{\left(\frac{1}{\gamma_{41}q^{n_0-1}}, \gamma_{31}q^{M_1}\right)_{N_1}} \overset{(7b)}{=\!=}$$

[Here the cast for (7b) is $\frac{r:=N_1, N:=n_0, x:=\frac{c'}{\gamma_{31}}}{r:=N_1, N:=n_0-1, x:=\gamma_{41}}$].

$$= \frac{\left(\frac{c'}{\gamma_{31}} q^{M_1+1}, \frac{c'}{\gamma_{41}} q^{M_1+1}\right)_{N_1}}{\left(\gamma_{41} q^{M_1}, \gamma_{31} q^{M_1}\right)_{N_1}} \cdot \left\{ \frac{(-1)^r \gamma_{41}^r q^{r(n_0-1)-\binom{r}{2}}}{(-1)^r \left(\frac{c'}{\gamma_{31}}\right)^r q^{rn_0-\binom{r}{2}}} = \left(\frac{\gamma_{31}\gamma_{41}}{c'q}\right)^r \right\}$$

Here we used the following cast for q-Pfaff (formula 9 of §2):
$a := c' q^{n_0+M_1}, b := \gamma_{51} q, c := \gamma_{31} q^{M_1}, N := N_1, n := n_1,$ therefore $\frac{ab}{cq^{N-1}} = \frac{\gamma_{51} c' q^{n_0+M_1+1}}{\gamma_{31} q^{M_1+n_0-M_1-1}} = \gamma_{41} q^{M_1}$, i.e. in order to apply the q-Pfaff formula, we must have $\boxed{\frac{\gamma_{51} c' q^2}{\gamma_{31}} = \gamma_{41}}$ So we can write

$$D_1 \cdot \sum_{n_1=0}^{N_1} G_1 B_1 = \frac{(\gamma_{11}, \gamma_{21})_{M_1}}{(\gamma_{31}, \gamma_{41})_{M_1}} (\gamma_{61} q)^{M_1} \cdot \frac{\left(\frac{c'}{\gamma_{31}} q^{M_1+1}, \frac{c'}{\gamma_{41}} q^{M_1+1}\right)_{N_1}}{\left(\gamma_{41} q^{M_1}, \gamma_{31} q^{M_1}\right)_{N_1}} \left(\frac{\gamma_{31}\gamma_{41}}{c'q}\right)^{N_1} .$$

Therefore if $\boxed{\gamma_{11} = \frac{c'q}{\gamma_{31}}, \gamma_{21} = \frac{c'q}{\gamma_{41}}, \gamma_{61} q = \frac{\gamma_{31}\gamma_{41}}{c'q}}$ we have

$$D_1 \cdot \sum_{n_1=0}^{N_1} G_1 B_1 = \frac{(\gamma_{11}, \gamma_{21})_{n_0}}{(\gamma_{31}, \gamma_{41})_{n_0}} \cdot (\gamma_{61} q)^{n_0}$$

Similarly, if the conditions

$\boxed{\gamma_{5i} = \gamma_{6i} = \frac{\gamma_{3i}\gamma_{4i}}{c' q^2}, \gamma_{1i} = \frac{c'q}{\gamma_{3i}}, \gamma_{2i} = \frac{c'q}{\gamma_{4i}}}$ for $i = 1, 2, \ldots, \ell - 2$ are satisfied,

then

$$D_i \cdot \sum_{n_i=0}^{N_i} G_i B_i = \frac{(\gamma_{1i}, \gamma_{2i})_{n_0}}{(\gamma_{3i}, \gamma_{4i})_{n_0}} \cdot (\gamma_{6i} q)^{n_0} = \partial Q\,(n_0, n_0)$$

Finally,

$$\sum_{n_{\ell-1}=0}^{N_{\ell-1}} P_c\left(n_{\ell-1}+n_\ell, n_\ell\right) B_{\ell-1} =$$

$$\frac{1}{(cq)_{2n_\ell}} \sum_{n_{\ell-1}=0}^{N_{\ell-1}} \frac{\left(c' q^{n_0+n_\ell}, \frac{1}{q^{N_\ell-1}}\right)_{n_{\ell-1}}}{\left(q, cq^{2n_\ell+1}\right)_{n_{\ell-1}}} q^{n_{\ell-1}}$$

$$\overset{(10)}{=\!=} \left[cast: N := N_{\ell-1},\ a := c' q^{n_0+n_\ell},\ c := cq^{2n_\ell+1}\right]$$

$$= \frac{1}{(cq)_{2n_\ell}} \frac{\left(\frac{c}{c'}\frac{1}{q^{n_0-n_\ell-1}}\right)_{N_{\ell-1}}}{\left(cq^{2n_\ell+1}\right)_{N_{\ell-1}}} \left(c' q^{n_0+n_\ell}\right)^{N_{\ell-1}} \overset{(7a)}{=\!=} \overset{x:=\frac{c'}{c}}{=\!=} \overset{r:=0}{=\!=}$$

$$= \frac{\left(\frac{c'}{c}\right)_{N_{\ell-1}}}{(-1)^{N_{\ell-1}}\left(\frac{c'}{c}\right)^{N_{\ell-1}} q^{\binom{N_{\ell-1}}{2}}} \cdot \frac{\left(c' q^{n_0+n_\ell}\right)^{N_{\ell-1}}}{(cq)_{n_0+n_\ell}} =$$

$$= \frac{\left(\frac{c'}{c}\right)_{n_0-n_\ell}}{(-1)^{n_0-n_\ell} q^{\binom{n_0-n_\ell}{2}}} \frac{\left(cq^{n_0+n_\ell}\right)^{n_0-n_\ell}}{(cq)_{n_0+n_\ell}}.$$

Since

$$B_\ell = \left(\frac{1}{q^{N_\ell}}, c' q^{n_0}\right)_{n_\ell} q^{n_\ell} \overset{(7b)}{=\!=} \frac{\left(q^{N_\ell-n_\ell+1}\right)_{n_\ell}}{(-1)^{n_\ell} q^{n_\ell N_\ell-\binom{n_\ell}{2}}} \cdot \left(c' q^{n_0}\right)_{n_\ell} q^{n_\ell},$$

we have

$$B_\ell \sum_{n_{\ell-1}=0}^{N_{\ell-1}} P_c B_{\ell-1} \overset{(4)}{=\!=} \left(c' q^{n_0}\right)_{n_\ell} q^{n_\ell} \frac{(q)_{n_0}}{(q)_{n_0-n_\ell}} \frac{\left(\frac{c'}{c}\right)_{n_0-n_\ell}}{(-1)^{n_0} q^{n_\ell+\binom{n_0}{2}}} \cdot \frac{\left(cq^{n_0+n_\ell}\right)^{n_0-n_\ell}}{(cq)_{n_0+n_\ell}} =$$

$$= (-1)^{n_0} q^{-\binom{n_0}{2}} \frac{\left(\frac{c'}{c}\right)_{n_0-n_\ell}}{(q)_{n_0-n_\ell}} \frac{(q)_{n_0}\left(c' q^{n_0}\right)_{n_\ell}}{(cq)_{n_0+n_\ell}} \cdot \left(cq^{n_0+n_\ell}\right)^{n_0-n_\ell}.$$

Therefore

$$B_0 B_\ell \sum_{n_{\ell-1}=0}^{N_{\ell-1}} P_c B_{\ell-1} = (-1)^{n_0} q^{\binom{n_0}{2}} \frac{\left(c', q\sqrt{c'}, -q\sqrt{c'}\right)_{n_0}}{\left(q, \sqrt{c'}, -\sqrt{c'}\right)_{n_0}} \cdot (\text{expression above}) =$$

$$\frac{\left(\frac{c'}{c}\right)_{n_0-n_\ell} (c')_{n_0+n_\ell}}{(q)_{n_0-n_\ell} (cq)_{n_0+n_\ell}} \frac{1-c'q^{2n_0}}{1-c'} \cdot \left(cq^{n_0+n_\ell}\right)^{n_0-n_\ell} .$$

∎

§4.3. Elementary Properties and Identities Used in the Proof

The formulas here can be found in [Gas90], [And76] or can be deduced easily.

Limits:

$$\lim_{a\to\infty} \frac{(a)_n}{a^n} = \begin{cases} (-1)^n q^{\binom{n}{2}}, & \text{if } if \\ 1, & \text{if } n \geqslant 1 \\ (-1)^n \frac{1}{q^{\binom{-n+1}{2}}}, & \text{if } n = 0. \end{cases} \tag{1}$$

$$\lim_{\gamma_1,\gamma_2\to\infty} \frac{(\gamma_1, \gamma_2)_n}{(\gamma_1 \cdot \gamma_2)^n} = \begin{cases} q^{n^2-n}, & \text{if } if \\ 1, & \text{if } n \geqslant 1 \\ \frac{1}{q^{n^2-n}}, & \text{if } n = 0. \end{cases} \tag{1a}$$

$$\lim_{a\to\infty} \frac{(a)_n}{\left(\frac{a}{x}\right)_n} = x^n \tag{2}$$

Hidden zeros:

Let $N \geq 1$.

$$if\ n \geq N+1\ then\ \left(\frac{1}{q^N}\right)_n = 0 \tag{3a}$$

$$if \ n \geq N \ then \ \frac{1}{(q^N)_{-n}} = 0 \tag{3b}$$

Identities:

$$\binom{N-r}{2} = \binom{N}{2} + \left[\binom{r}{2} - N \cdot r\right] + r \tag{4}$$

$$\binom{r-N}{2} = \binom{N+1}{2} + \left[\binom{r}{2} - N \cdot r\right] \tag{4a}$$

In the following, let M, N, n, r be ≥ 0.

$$\frac{1-cq^{2n}}{1-c} = \frac{(q\sqrt{c}, -q\sqrt{c})_n}{(\sqrt{c}, -\sqrt{c})_n} \tag{5}$$

$$(a)_{n+r} = (a)_r \cdot (aq^r)_n \tag{6}$$

$$\frac{1}{(x)_{N-r}} = \frac{1}{(x)_N} \cdot \left(xq^{N-r}\right)_r \tag{6a}$$

$$\frac{1}{(q)_{N-r}} = \frac{1}{(q)_N} \cdot \left(q^{N-r+1}\right)_r \tag{6b[x=q]}$$

$$\left(\frac{x}{q^N}\right)_{N+M} = \frac{1}{(x)_{-N}} \cdot (x)_M \tag{6c}$$

$$(xq^r)_N = (-1)^N \, x^N q^{N \cdot r + \binom{N}{2}} \cdot \left(\frac{1}{xq^{r+N-1}}\right)_N \tag{7a}$$

$$\left(xq^{N-r+1}\right)_r = (-1)^r \, x^r q^{N \cdot r - \binom{r}{2}} \cdot \left(\frac{1}{xq^N}\right)_r \tag{7b}$$

Identities (6a) and (7b) give

$$\boxed{\frac{1}{(x)_{N-r}} = \frac{1}{(x)_N} \cdot (-1)^r \, x^r q^{N \cdot r - \binom{r+1}{2}} \cdot \left(\frac{1}{xq^{N-1}}\right)_r} \tag{8}$$

$$\frac{1}{(q)_{N-r}} = \frac{1}{(q)_N} \cdot (-1)^r \, x^r q^{N \cdot r - \binom{r}{2}} \cdot \left(\frac{1}{q^N}\right)_r \qquad \text{(8a [x=q])}$$

$$\frac{1}{(x)_{-r}} = (-1)^r \, x^r q^{-\binom{r+1}{2}} \cdot \left(\frac{q}{x}\right)_r . \qquad \text{(8b [N=0])}$$

The following formula *[q-Pfaff-Saalschütz]* is a real theorem:

$$\sum_{n=0}^{N} \frac{\left(a, b, \dfrac{1}{q^N}\right)_n}{\left(q, c, \dfrac{ab}{cq^{N-1}}\right)_n} q^n = \frac{\left(\dfrac{c}{a}, \dfrac{c}{b}\right)_N}{\left(c, \dfrac{c}{ab}\right)_N}. \tag{9}$$

If we take here $b \to 0$ then, by (2) we obtain the *q-Chu-Vandermonde* formula:

$$\sum_{n=0}^{N} \frac{\left(a, \dfrac{1}{q^N}\right)_n}{(q, c)_n} q^n = \frac{\left(\dfrac{c}{a}\right)_N}{(c)_N} a^N . \tag{10}$$

CHAPTER 5

ORBITS OF ITERATED AFFINE TRANSFORMATIONS

§5.1. Introduction

Transformations $T(\underline{x}) := A\underline{x}+\underline{b}$, where A is a matrix and $\underline{x}$ is a vector are called affine transformations. (In this paper expressions like $T := A$ or $A =: T$ should be understood as 'T is defined by A'. In addition, vectors are underlined and matrices are capitalized). Affine transformations are essential in theoretical mathematics, such as partition theory, in robotics [Har92], and in image compression [Bar93]. For instance, images with self-similarities can be encoded by storing affine transformations with their probabilities and can be regained by applying the transformations in a random way to any point in the image plane, leading to quick approximation of the original image.

We present two main results:

First, we develop a general formula to compute the orbit produced by iterating an affine transformation. To achieve this, the main idea is to compute certain projections of the orbit. Second, we will use novel matrix products to reduce the number of array multiplications. One of these products is the generalized matrix product, introduced by G. Ritter [Rit91].

The main body of the paper is divided into seven sections:

In §2. we give a solution (Proposition 1) using the Jordan normal form of A. We will give an alternative solution in §2- §5.

In §3. we present a method to find left eigenvectors by solving a system of quadratic equations. We also show an important iterative step.

In §4. we introduce some essential tools from approximation theory and obtain three lemmas for later use (Lemma 1,2,3).

In §5. we give an algorithm to obtain formulas for certain projections of smaller and smaller tails of the orbit (Theorem 1). This leads to a method to compute the orbit itself, even in the worst case, when all of our matrices with shrinking dimensions are defective (Corollary 1).

In §6., after some preparation (Lemma 4, Proposition 2), we get a formula for the orbit, provided that at least one of the shrinking matrices is non-defective (Theorem 2).

In §7. we introduce novel matrix products to reduce the computational costs in an array computer.

In §8. we highlight the two-dimensional case with formulas in terms of both the traditional and the new matrix products and show how to obtain the usual formulas for recurrence sequences found in number theory books by specialization.

§5.2. The Solution Using the Jordan Normal Form

Our investigation starts with the following observation:

Let A be a $d \times d$ dimensional matrix and let us suppose that its inverse, denoted either by A^{-1} or by $\frac{1}{A}$, exists. If T is an affine transformation, $T(\underline{x}) := A\underline{x} + \underline{b}$, then its inverse, T^{-1} satisfies $T^{-1}(\underline{x}) = A^{-1}\underline{x} - A^{-1}\underline{b}$. Using induction, we obtain the following result:

PROPOSITION 1. *For any, positive or negative integer n the identity*

$$T^n(\underline{x}) = A^n\underline{x} + (A)_n\underline{b}, \tag{1}$$

holds, where T^n and A^n are the $|n|^{th}$ iterative and multiplicative powers of $T^{sign\, n}$ and $A^{sign\, n}$ respectively, and

$$(A)_n := (0; A)_n := \begin{cases} 1 + A + \ldots + A^{n-1} & if\, n > 0 \\ 0 & if\, n = 0 \\ -\left(A^{-1} + A^{-2} + \ldots + A^{n}\right), & if\, n < 0. \end{cases} \tag{2}$$

If there is a sufficient number (i.e. d) of right or left eigenvectors for A, then A can be decomposed as $L^{-1}DL$, where D is a diagonal matrix containing the eigenvalues μ_j and L is a matrix storing the eigenvectors. Therefore, we have that

$$T^n(\underline{x}) = L^{-1}\, diag\,[\,\mu_1^n, \mu_2^n, \ldots, \mu_d^n\,]\; L\underline{x} + L^{-1}\, diag\,[\,(\mu_1)_n, (\mu_2)_n, \ldots, (\mu_d)_n\,]\; L\underline{b} \tag{3}$$

Let us remark that in definition (2), not only can A be a matrix, but a number as well. In addition, we can extend the definition of $(0; A)_n$ by setting $(x; A)_n := nx + (0; A)_n$.

From (3) we can derive some important properties of affine transformations:

First, T is cyclic with period n iff all μ_j's are n^{th} roots of unity, i.e. $\mu_j = \omega^{\alpha_j}$, where $\omega := exp(\frac{2\pi i}{n})$ and α_j is an integer.

Second, $\mu_j^n \to 0$ and $(\mu_j)_n \to \frac{1}{1-\mu_j}$ iff either $|\mu_j| < 1$ as $n \to \infty$ or $|\mu_j| > 1$ as $n \to -\infty$.

If we do not have enough eigenvectors, then the matrix is called *defective*. In this case we can still use the Jordan form J of A to write A as $C^{-1}JC$ to get

PROPOSITION 1'. *For any, positive or negative integer n the identity*

$$T^n(\underline{x}) = C^{-1}\left(J^n C\underline{x} + (J)_n C\underline{b}\right), \tag{3'}$$

holds.

Let $\begin{bmatrix} \mu & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \mu \end{bmatrix}$ be a $(\alpha \times \alpha)$-dimensinal diagonal block of J. Using the general binomial theorem, write $(\mu + 1)^n = \sum_{k=0}^{\infty} c_k$, where $c_k := \binom{n}{k}\mu^{n-k}$. Recall that if $n \geq 0$ and $k > n$, then $\binom{n}{k} = 0$; on the other hand, if $n < 0$, then $\binom{n}{k} = (-1)^k \binom{-n+k-1}{k}$.

The corresponding block of J^n is the upper triangular matrix

$$\begin{bmatrix} c_0 & c_1 & \cdots & \cdots & c_{g-1} \\ & c_0 & c_1 & \cdots & c_{g-2} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & c_1 \\ & & & & c_0 \end{bmatrix}.$$

How can we obtain the transformation matrix C with good precision? Let us outline a possible solution. Use Danilevski's method (see [Ham70]) to obtain matrices B and D at the cost of $O(d^3)$ scalar multiplication and divisions, such that $A = D^{-1}BD$ and B is in the 'companion matrix' form, i.e.

$$B = \begin{bmatrix} b_1 & b_2 & \dots & b_{d-1} & b_d \\ 1 & & & & 0 \\ & 1 & & & \vdots \\ & & \ddots & & \vdots \\ & & & 1 & 0 \end{bmatrix},$$

where the unmarked entries are zeros. The characteristic polynomial of B equals that of A, i.e.

$t^d - \left(b_1 t^{d-1} + b_2 t^{d-2} + b_{d-1} t + b_d\right) = (t-\mu_1)^{\alpha_1} \cdot \ldots \cdot (t-\mu_r)^{\alpha_r}$.

Let $V_s := \left[\binom{d-j}{k-1}\mu_s^{d-k-j+1}\right]$, where the row index j runs from 1 to d, while column index k runs from 1 to α_s.

The *generalized Vandermonde* matrix, $V := [V_1, V_2, \ldots, V_r]$, or rather its determinant, seems to go back to L. Schendel (1891), quoted by [Muil1, Vol. IV, page 179]. It has a non-zero determinant: *det* $V = \prod_{k>j}\left(\mu_k - \mu_j\right)^{\alpha_j \alpha_k}$. (The case $\alpha_1 = \ldots = \alpha_r = 1$ is the usual Vandermonde matrix). In addition, $B = VJV^{-1}$ (see [Ait56, page 136]), which gives $A = \left(V^{-1}D\right)^{-1} J \left(V^{-1}D\right)$. In the following sections, we give alternative formulas for $T^n(\underline{x})$, which do not rely on the Jordan decomposition.

§5.3. Projection of the Orbit, First Descent

In this section we give a method how to obtain the left eigenvectors of a matrix first, and then its eigenvalues. The method itself is not needed in the rest of the paper, except one property: the first coordinate of every eigenvector is 1.

But first, let us observe that an affine transformation T can always be written in a matrix form:

$$T(\underline{x}) := \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_d & \underline{b} \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{x} \\ 1 \end{bmatrix} := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,d} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2,d} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \cdots & a_{d,d} & b_d \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \\ 1 \end{bmatrix}$$

This has the advantage that iteration of T can be regarded as a matrix multiplication. Let us examine the orbit generated by T: let

$$\begin{bmatrix} \underline{x}(n) \\ 1 \end{bmatrix} := \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_d & \underline{b} \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{x}(n-1) \\ 1 \end{bmatrix}$$

Let $\underline{\lambda} := [\lambda_1, \lambda_2, \ldots, \lambda_d]$ be any row vector. We have $\underline{\lambda} \cdot \underline{x}(n) = (\underline{\lambda a_1}) \cdot x_1(n-1) + \ldots + (\underline{\lambda a_d}) \cdot x_d(n-1) + \underline{\lambda b}$. Fix $\lambda_1 = 1$ and let $\underline{\lambda}$ be a solution of the system of quadratic equations

$$\underline{\lambda a_1} = \frac{1}{\lambda_2} \cdot \underline{\lambda a_2} = \ldots = \frac{1}{\lambda_d} \cdot \underline{\lambda a_d} \tag{4}$$

We have $(d-1)$ quadratic equations of the $(d-1)$ unknowns $\lambda_2, \lambda_3, \ldots, \lambda_d$. If we take $\mu := \underline{\lambda a_1}$, we can see that solving (4) is equivalent to solving the left eigenvector problem $\underline{\lambda}\mu = \underline{\lambda}A$. Therefore, we can have $1, 2, \ldots, d$ solutions $\underline{\lambda}$.

If $d = 2$, we have two solutions iff $(a_{22} - a_{11})^2 + 4a_{12}a_{21} \neq 0$.

If $d = 3$, the solutions for $x = \lambda_2$ of (4) are the roots of the cubic equation $g_0(GA)x^3 + g_1(GA)x^2 + g_1(A)x + g_0(A) = 0$, where
$g_0(a) := a_{21}^2 a_{13} - a_{23}^2 a_{31} + a_{21}a_{23}(a_{33} - a_{11})$,
$g_1(a) := a_{11}^2 a_{23} - a_{23}^2 a_{32} + a_{33}a_{23}(a_{22} - a_{11}) - a_{23}(a_{11} + a_{22} + a_{12}a_{21}) + 2a_{13}(a_{23}a_{31} + a_{21}a_{22})$, and G is the transformation that swaps the pairs
$(a_{12}, a_{21}), (a_{11}, a_{22}), (a_{13}, a_{23}), (a_{31}, a_{32})$ and does not change a_{33}.[1]

For any solution $\underline{\lambda}$ of (4) we have

$$\underline{\lambda} \cdot \underline{x}(n) = \mu(x_1(n-1) + \lambda_2 \cdot x_2(n-1) + \ldots + \lambda_d \cdot x_d(n-1)) + \underline{\lambda} \cdot \underline{b} = \mu \cdot \underline{\lambda x}(n-1) + \underline{\lambda} \cdot \underline{b}. \tag{4}$$

Iterate this for $n-1, n-2$,etc. to get

$$\underline{\lambda} \cdot \underline{x}(n) = \mu^n \cdot \underline{\lambda x}(0) + (\mu)_n \cdot \underline{\lambda b} =: R_n. \tag{5}$$

An important note is required here. If matrix A is not constant, but varies, i.e. $A = A(n)$, and if we require only that the eigenvector λ stay fixed, but eigenvalues

[1] I am indebted to Dr. Ralph Selfridge, who produced this polynomial at my request, using his APL-based quadratic system solving package.

$\mu(n)$ can change (for example, matrices of continued fractions are like that), then the iterative step above can be carried through. If we redefine μ^n and $(\mu)_n$ by setting $\mu^n := \mu(1) \cdot \mu(2) \cdot \ldots \cdot \mu(n)$ and
$(\mu)_n := \mu(1) \cdot \ldots \cdot \mu(n-1) + \mu(2) \cdot \ldots \cdot \mu(n-1) + \ldots + \mu(n-1) + 1$, then we obtain (5) again.

In this section, we have obtained a formula for a projection of the orbit, $\underline{\lambda} \cdot \underline{x}(n)$. Of course, our aim is to get a formula for $\underline{x}(n)$. We achieve this in §4, but we have to build some tools first.

§5.4. Divided Differences. Some Notation

The main goal of this section is to prove Lemma 3. below. Let J be an index set, let x_j, $j \in J$ be all distinct numbers, $k \in J$, $\varpi_k(x) := \prod_{j \neq k}(x - x_j)$ and $\varpi(x) := (x - x_k)\varpi_k(x) = \prod_{j \in J}(x - x_j)$.

LEMMA 1. *[Abel] If $|J| \geq 2$, then*

$$\sum_{j \in J} \frac{1}{\varpi'(x_j)} = 0.$$

PROOF: The *Lagrange* interpolation formula gives us that if $P(x)$ is any polynomial of degree $\leq |J| - 2$ then

$$P(x) = \sum_{j \neq k} P(x_j) \frac{\varpi_k(x)}{(x - x_j)\varpi_k'(x_j)}$$

. Take $P(x) \equiv 1$, divide by $\varpi(x)$ and move everything on one side to obtain

$$\frac{1}{\varpi_k(x)} + \sum_{j \neq k} \frac{1}{(x_j - x)\varpi_k'(x_j)} = 0.$$

Substitute $x = x_k$ and consider

$$\varpi'(x) = (x - x_k)\varpi_k'(x) + \varpi_k(x) = \begin{cases} \varpi_k(x_k) & \text{if } x = x_k \\ (x_j - x_k)\varpi_k'(x_j) & \text{if } x = x_j, j \neq k \end{cases}$$

in order to obtain the result. ∎

An easy corollary is

LEMMA 2. *We have the identity*

$$\sum_{j \neq k} \frac{\frac{f(x_j) - f(x_k)}{x_j - x_k}}{\varpi_k'(x_j)} = \sum_j \frac{f(x_j)}{\varpi'(x_j)}$$

PROOF: From Lemma 1 we have

$$\frac{1}{\varpi'(x_k)} = -\sum_{j \neq k} \frac{1}{\varpi'(x_j)} = -\sum_{j \neq k} \frac{1}{(x_j - x_k)\varpi_k'(x_j)},$$

which implies

$$\begin{aligned} \sum_{j \neq k} \frac{f(x_j) - f(x_k)}{(x_j - x_k)\varpi_k'(x_j)} &= \sum_{j \neq k} \frac{f(x_j)}{\varpi'(x_j)} - f(x_k) \cdot \sum_{j \neq k} \frac{1}{(x_j - x_k)\varpi_k'(x_j)} \\ &= \sum_{j \neq k} \frac{f(x_j)}{\varpi'(x_j)} + f(x_k) \cdot \frac{1}{\varpi'(x_k)} = \sum_j \frac{f(x_j)}{\varpi'(x_j)} \end{aligned}$$ ∎

Suppose we are given a sequence of numbers: $x_0, x_1, \ldots$. The *divided difference* is defined recursively in approximation theory:

$$[x_0]f := f(x_0), \ [x_0, x_1 \ldots, x_k, x_{k+1}]f = \frac{[x_1, \ldots, x_k, x_{k+1}]f - [x_0, x_1, \ldots, x_k]f}{x_{k+1} - x_0}$$

We know that [Gel71]

$$[x_0, x_1, \ldots, x_k]f = \sum_{j=0}^{k} \frac{f(x_j)}{\varpi'(x_j)}, \text{ where } \varpi(x) := \prod_{j=0}^{k}(x - x_j),$$

and

$$\lim_{\text{all } j<k:\ x_j \to x_k} [x_0, x_1, \ldots, x_{k-1}, x_k]f = \frac{f^{(k)}(x_k)}{k!}. \tag{6}$$

We will need the divided differences of the functions $f_n(x) := x^n$ and $f_n(x) := (x)_n$. The latter was defined in (2).

Lemma 3. *For both functions we have*

$$\sum_{m=0}^{n-1} [x_0, \ldots, x_{k-1}]f_{n-1-m} \cdot x_k^m = [x_0, \ldots, x_{k-1}, x_k]f_n$$

Proof: Let $k = 1$, $a := x(0)$, $b := x(1)$. For $f(x) := x^n$, the statement is just the identity $\sum a^{n-1-m}b^m = (a^n - b^n)/(a - b)$. To prove

$$\sum_{m=0}^{n-1} (a)_{n-1-m} b^m = \frac{(a)_n - (b)_n}{a - b} \tag{7}$$

regard the matrix

$$\begin{bmatrix} a^{n-2} & a^{n-3} & \ldots & a^2 & a & 1 \\ a^{n-3}b & a^{n-4}b & \ldots & ab & b & \\ \vdots & \vdots & & & & \\ ab^{n-3} & b^{n-3} & & & & \\ b^{n-2} & & & & & \end{bmatrix}.$$

On the one hand, the sum of the entries, if we first add the rows together is

$$\sum_{r=1}^{n-1} (a)_{n-r} b^{r-1} = \sum_{r=1}^{n} \text{ the same } = \sum_{m=0}^{n-1} (a)_{n-m-1} b^m.$$

On the other hand, if we obtain the column sums first, we have

$$\sum_{c=1}^{n-1} \frac{a^{n-c} - b^{n-c}}{a-b} = \sum_{c=1}^{n} \text{the same} = \frac{(a)_n - (b)_n}{a-b}.$$

If $k > 1$, the left hand side of the claim is equal to

$$\sum_{m=0}^{n-1} \sum_{j=0}^{k-1} \frac{f_{n-1-m}(x_j)}{\varpi'_k(x_j)} x_k^m =^{(\text{by }(7))} \sum_{j=0}^{k-1} \frac{f_n(x_j) - f_n(x_k)}{(x_j - x_k)\varpi'_k(x_j)} =$$

$$=^{(\text{by Lemma 2})} \sum_{j=0}^{k} \frac{f_n(x_j)}{\varpi'(x_j)} = RHS. \blacksquare$$

Lemma 3. will be used in the proof of Theorem 1. in the next section. Let us introduce the following two abbreviations:

$[x_0, x_1, \ldots, x_k]^n := [x_0, x_1, \ldots, x_k] f_n$ where $f_n(x) := x^n$ and

$[x_0, x_1, \ldots, x_k]_n := [x_0, x_1, \ldots, x_k] f_n$ where $f_n(x) := (x)_n$.

We close this section by introducing some notation, which will be used in the next section. If $\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}$ is any vector, let $\underline{y}_{>k}$ be the tail end of it, i.e. $\underline{y}_{>k} := \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ \vdots \\ y_d \end{bmatrix}$. Similarly, if $A = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{dd} \end{bmatrix}$ then $A_{>k}$ will stand for the submatrix $\begin{bmatrix} a_{k+1,k+1} & & \\ & \ddots & \\ & & a_{dd} \end{bmatrix}$. For instance, $\underline{a}_{>k,\ell} = \begin{bmatrix} a_{k+1,\ell} \\ a_{k+2,\ell} \\ \vdots \\ a_{d,\ell} \end{bmatrix}$.

Suppose we have a sequence of matrices $A(k) = [\underline{a}_{>k,k+1}(k), \ldots, \underline{a}_{>k,d}(k)]$. Here $dim\ A(k) = (d-k) \times (d-k)$. Suppose also that vectors $\underline{\lambda}(k)$ and scalars $\mu(k)$ are given too. Define

$\rho(k,k) := 1,$

$\rho(k_1,k_2) := \underline{\lambda}(k_1) \cdot \underline{a}_{>k_1,k_2+1}(k_2)$ if $k_1 > k_2$,

$$\rho(k_1,\ldots,k_s) := \prod_{j=1}^{s-1} \rho(k_j,k_{j+1}) \quad if\ k_1 \geq k_2 \geq \ldots \geq k_s.$$

There is a last series of definitions left:

$[k \searrow k]f := [\mu_k]f$

$[k \searrow k-1]f := \rho(k,k-1) \cdot [\mu_k,\mu_{k-1}]f$

$[k \searrow k-2]f := \rho(k,k-1,k-2) \cdot [\mu_k,\mu_{k-1},\mu_{k-2}]f + \rho(k,k-2) \cdot [\mu_k,\mu_{k-2}]f$

$$[k \searrow k-3]f := \rho(k,k-1,k-2,k-3) \cdot [\mu(k),\mu(k-1),\mu(k-2),\mu(k-3)]f+$$
$$\rho(k,k-1,k-3) \cdot [\mu(k),\mu(k-1),\mu(k-3)]f + \rho(k,k-2,k-3) \cdot [\mu(k),\mu(k-2),\mu(k-3)]f+$$
$$\rho(k,k-3) \cdot [\mu(k),\mu(k-3)]f$$

In general, let $[k \searrow j] := \sum_{\pi(k-j)} \rho(k,\ldots,j) \cdot [\mu(k),\ldots,\mu(j)]f$, where the summation extends to all partitions of $(k-j)$ and the order of parts counts.

§5.5. Projection of the Orbit, Full Descent

At the end of §2, we established the identity

$$\underline{\lambda} \cdot \underline{x}(n) = \mu^n \cdot \underline{\lambda x}(0) + (\mu)_n \cdot \underline{\lambda b} =: R_n.$$

Therefore a projection of the orbit $\underline{x}(n)$ can be computed. The algorithm below gives a sequence of square matrices with sizes $d, d-1, \ldots, 1$. At each step we take a left

eigenvector and an eigenvalue, and compute a new projection of the tail of the orbit. At step $d-1$ (i.e. at size 1) we obtain a formula for $x_d(n)$ itself, due to the fact that the first coordinate of each eigenvector is normalized to be 1. The rest of the coordinates, $x_j(n)$, can be found one after the other by stepping back.

THEOREM 1. *Let* $A(0) := A, \underline{\lambda}(0) := \underline{\lambda}, \mu(0) := \mu$ *from §2.*

If

$$A(k) := A_{>k}(k-1) - \underline{a}_{>k,k}(k-1) \cdot \underline{\lambda}_{>k}(k-1), \tag{8}$$

$\underline{\lambda}(k) = [1, \lambda_{k+2}(k), \ldots, \lambda_d(k)]$ *and* $\mu(k)$ *satisfy*

$$\underline{\lambda}(k)\mu(k) = \underline{\lambda}(k)A(k), \tag{9}$$

then

$$\boxed{\underline{\lambda}(k) \cdot \underline{x}_{>k}(n) = \sum_{j=0}^{k} [k \searrow j]^n \cdot \underline{\lambda}(j) \cdot \underline{x}_{>j}(0) + \sum_{j=0}^{k} [k \searrow j]_n \cdot \underline{\lambda}(j) \cdot \underline{b}_{>j} =: R_n(k).} \tag{10}$$

In addition, we have

$$x_k(n) = -\underline{\lambda}_{>k}(k-1) \cdot \underline{x}_{>k}(n) + R_n(k-1). \tag{11}$$

and

$$\underline{x}_{>k}(n) = A(k) \cdot \underline{x}_{>k}(n-1) + \underline{b}_{>k} + \sum_{j=0}^{k-1} \underline{a}_{>k,j+1}(j) \cdot R_{n-1}(j). \tag{12}$$

NOTE. *As in §2,* $\underline{\lambda}(k)$ *and* $\mu(k)$ *can be obtained by solving the quadratic system of equations*

$$\underline{\lambda}(k)\underline{a}_{k+1}(k) = \frac{1}{\lambda_{k+2}(k)} \cdot \underline{\lambda}(k)\underline{a}_{k+2}(k) = \ldots = \frac{1}{\lambda_d(k)} \cdot \underline{\lambda}(k)\underline{a}_d(k),$$

and by setting

$$\mu(k) := \underline{\lambda}(k)\underline{a}_{k+1}(k).$$

COROLLARY 1. *Since $\underline{\lambda}(d-1) = [1]$, we have*

$$x_d(n) = \underline{x}_{>d-1}(n) = R_n(d-1).$$

Therefore we can obtain $x_{d-1}(n), x_{d-2}(n), \ldots, x_1(n)$ one after the other, using (11).

PROOF: We use induction on k. The $k=0$ case was discussed in §2 Let $k \geq 1$ and suppose the theorem holds for $k-1$. Since $\underline{\lambda}_k(k-1) = 1$, we have (11). Because (12) is supposed to be true for $k-1$, we have

$$\begin{aligned}
&\underline{x}_{>k}(n) = \\
&\underline{a}_{>k,k}(k-1) \cdot x_k(n-1) + A_{>k}(k-1) \cdot \underline{x}_{>k}(n-1) + \underline{b}_{>k} + \sum_{j=0}^{k-2} \underline{a}_{>k,j+1}(j) R_{n-1}(j) \\
&= \underline{a}_{>k,k}(k-1) \cdot (-\underline{\lambda}_{>k}(k-1) \cdot \underline{x}_{>k}(n) + R_n(k-1)) + A_{>k}(n-1) \cdot \underline{x}_{>k}(n-1) + \underline{b}_{>k} + \\
&\qquad + \sum_{j=0}^{k-2} \underline{a}_{>k,j+1}(j) R_{n-1}(j) \\
&= A(k)\underline{x}_{>k}(k-1) + a_{>k,k}(k-1) \cdot R_{n-1}(k-1) + \underline{b}_{>k} + \sum_{j=0}^{k-2} \underline{a}_{>k,j+1}(j) R_{n-1}(j),
\end{aligned}$$

which proves (12) for k. If $\underline{\lambda}(k)$ and $\mu(k)$ are chosen by (9), we have, by the definition of ρ,

$$\begin{aligned}
&\underline{\lambda}(k)\underline{x}_{>k}(n) = \\
&\quad \underline{\lambda}(k)A(k)\underline{x}_{>k}(n-1) + \underline{\lambda}(k) \cdot \left(\underline{b}_{>k} + \sum_{j=0}^{k-1} \underline{a}_{>k,j+1}(j) \cdot R_{n-1}(j) \right) \\
&\qquad = \mu(k) \cdot \underline{\lambda}(k)\underline{x}_{>k}(n-1) + \underline{\lambda}(k)\underline{b}_{>k} + \sum_{j=0}^{k-1} \rho(k,j) R_{n-1}(j).
\end{aligned}$$

Iterate this to get

$$\underline{\lambda}(k)\underline{x}_{>k}(n) = [\mu(k)]^n\underline{\lambda}(k)\underline{x}_{>k}(0) + [\mu(k)]_n\underline{\lambda}(k)\underline{b}_{>k} + R_n^*(k),$$

where

$$R_n^*(k) := \sum_{m=0}^{n-1}[\mu(k)]^m \cdot \sum_{j=0}^{k-1}\rho(k,j)R_{n-1-m}(j).$$

But

$$R_n^*(k) = \sum_{m=0}^{n-1}\sum_{j=0}^{k-1}\sum_{i=0}^{j}\sum_{\pi(j-i)} (\rho(k,j)\rho(j,\ldots,i)[\mu(k)]^m[\mu(j),\ldots,\mu(i)]^{n-1-m}\underline{\lambda}(i)\underline{x}_{>i}(0) + \rho(k,j)\rho(j,\ldots,i)[\mu(k)]^m[\mu(j),\ldots,\mu(i)]_{n-1-m}\underline{\lambda}(i)\underline{b}_{>i})$$

$$=^{(by Lemma 3.)}$$

$$= \sum_{j=0}^{k-1}\sum_{i=0}^{j}\sum_{\pi(j-i)} \rho(k,j,\ldots,i)([\mu(k),\mu(j),\ldots,\mu(i)]^n\underline{\lambda}(i)\underline{x}_{>i}(0) + [\mu(k),\mu(j),\ldots,\mu(i)]_n\underline{\lambda}(i)\underline{b}_{>i}),$$

thus

$$R_n^*(k) = \sum_{i=0}^{k-1}[k \searrow i]^n\underline{\lambda}(i)\underline{x}_{>i}(0) + [k \searrow i]_n\underline{\lambda}(i)\underline{b}_{>i} \quad \blacksquare. \tag{13}$$

In the next section we will show how to reduce the number of steps in the algorithm of Theorem 1.

§5.6. Finding the Orbit by Shortcutting the Descent

In this section we show how to find the orbit if we experience that there is enough (i.e. $g := d - k$) eigenvectors at step k. (If it never happens, we can still use Corollary 1. from the previous section to get the orbit).

In the next two statements let $D := \begin{bmatrix} D_{11} & \dots & D_{1g} \\ \vdots & & \vdots \\ D_{g1} & \dots & D_{gg} \end{bmatrix}$ and

$$C := \begin{bmatrix} \underline{c}(1) \\ \vdots \\ \underline{c}(g) \end{bmatrix} := \begin{bmatrix} c_1(1) & \dots & c_g(1) \\ \vdots & & \vdots \\ c_1(g) & \dots & c_g(g) \end{bmatrix} \tag{14}$$

.

LEMMA 4. *If*

$$r(k) = \sum_{j=1}^{t} M_j(k)\underline{c}(k)\underline{y}_j + R = [\underline{c}(k)\cdot\underline{y}_1, \dots, \underline{c}(k)\cdot\underline{y}_t] \begin{bmatrix} M_1(k) \\ \vdots \\ M_t(k) \end{bmatrix} + R \tag{15}$$

then

$$\begin{bmatrix} \sum_{k=1}^{g} D_{1k}r(k) \\ \sum_{k=1}^{g} D_{2k}r(k) \\ \vdots \\ \sum_{k=1}^{g} D_{gk}r(k) \end{bmatrix} = \sum_{j=1}^{t} \left(D \cdot diag[M_j(1), M_j(2), \dots, M_j(g)] \cdot C\right) \underline{y}_j + R \cdot D \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

PROPOSITION 2. *The solution of the system of linear equations*

$c_1(1)x_1 + \dots + c_g(1)x_g = r(1) + R$

$c_1(2)x_1 + \dots + c_g(2)x_g = r(2) + R$

$\vdots$

$c_1(g)x_1 + \ldots + c_g(g)x_g = r(g) + R,$

where $r(k)$ *is in the form described by (15),* R *is a constant and* $\det C \neq 0$ *is*

$$\underline{x} = \sum_{j=1}^{t} \left(C^{-1} \cdot diag[M_j(1), M_j(2), \ldots, M_j(g)] \cdot C \right) \underline{y}_j + R \cdot C^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

PROOF: Let $D_{ji} := (-1)^{i+j} \det[\text{ row } i, \text{ column } j \text{ removed from } C]$.
By Cramér's rule, $x_j = \frac{(r(1)+R)\cdot D_{j1} + (r(2)+R)\cdot D_{j2} + \ldots + (r(g)+R)\cdot D_{jg}}{\det C}$, therefore

$$\underline{x} = \frac{1}{\det C} \begin{bmatrix} \sum_{k=1}^{g} D_{1k} r(k) \\ \sum_{k=1}^{g} D_{2k} r(k) \\ \vdots \\ \sum_{k=1}^{g} D_{gk} r(k) \end{bmatrix},$$

and since $C^{-1} = \frac{D}{\det C}$, the proof is finished by using the previous lemma. ∎

Let us go back to step k of the descent described in the previous paragraph. Let $\underline{c} := \underline{\lambda}(k)$ and

$$M := \mu(k) = \underline{c} \cdot \underline{a}_{k+1}(k) \tag{16}$$

THEOREM 2. *If we have enough,* $g := d - k$ *eigenvectors* $\{\underline{c}(j), j = 1, 2, \ldots, g\}$ *in*

step k of the descent, the corresponding eigenvalues $M(j)$ are given by (16), matrix C is defined by (14) then

$$
\begin{aligned}
\underline{x}_{>k}(n) &= \\
&= \left(C^{-1}\cdot diag[(M(1))^n,\dots,(M(g))^n]\cdot C\right)\underline{x}_{>k}(0) \\
&\quad + \left(C^{-1}\cdot diag[(M(1))_n,\dots,(M(g))_n]\cdot C\right)\underline{b}_{>k} \\
&\qquad + R_n^*(k)\cdot C^{-1}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix},
\end{aligned}
$$

where $R_n^*(k)$ is given by (13). The rest of the coordinates of the orbit, $\{x_k(n), x_{k-1}(n),\dots,x_1(n)\}$ can be obtained using (11) in §4.

PROOF: $R_n(k)$ in Theorem 1. is in the form (15):

$$R_n(k) = M^n \underline{cx}_{>k}(0) + (M)_n \underline{cb}_{>k} + R_n^*(k).$$

Use Proposition 2. with $t := 2$, $\underline{y}_1 := x_{>k}(0)$, $\underline{y}_2 := \underline{b}_{>k}$, $M_1(k) := (M(k))^n$ $M_2(k) := (M(k))_n$ and $R := R_n^*(k)$. ∎

§5.7. Computational Costs. Different Matrix Products

In this section, we would like to rewrite expressions like the one in Lemma 4.,

$$D\sum_{j=1}^{t}\left(diag[M_j(1), M_j(2),\dots,M_j(g)]\cdot C\right)\underline{y}_j. \tag{17}$$

Computing T^n in the straightforward way requires $(n-1)$ matrix multiplications. The formulas in the Theorems are special cases with $t = 2$ of (17), which invokes

$t+1 = 3$ array multiplications only, plus some set-up costs incurred by the algorithm in Theorem 1 and the storage of vectors $C\underline{y}_j$ and matrix D. We can rewrite (17) in terms of three different matrix products, $\cdot_1$, $\cdot_2$, $\cdot_3$, and reduce the number of array multiplications needed for iteration to 2. Expression (17) will take the form of

$$((D \cdot_1 C) \cdot_2 \underline{M}) \cdot_3 Y, \tag{18}$$

and here $D \cdot_1 C$ and Y can be stored.

The total number of scalar multiplications using (18) however, is larger than in (17): $g \cdot tg^2 + g \cdot tg$ vs. $tg + g^2$ (plus $(n-1)tg$ to compute the μ powers), but both of them are superior to the $[(g^2+g)g]^{n-1}$ (or, by Strassen's algorithm, $O(g^{log_2 7 \cdot (n-1)})$ scalar multiplications used in the straightforward method.

The definitions are the following:

$$[D_{j1}, \ldots, D_{jg}] \cdot_1 \begin{bmatrix} c_1(1) & \ldots & c_g(1) \\ \vdots & & \vdots \\ c_1(g) & \ldots & c_g(g) \end{bmatrix} :=$$

$$= [D_{j1}C_1(1), \ldots, D_{jg}C_1(g) | D_{j1}C_2(1), \ldots, D_{jg}C_2(g) | \ldots | D_{j1}C_g(1), \ldots, D_{jg}C_g(g)].$$

This is extended by $\begin{bmatrix} \underline{D}_j \\ \underline{D}_k \end{bmatrix} \cdot_1 C := \begin{bmatrix} \underline{D}_j \cdot_1 C \\ \underline{D}_k \cdot_1 C \end{bmatrix}$, etc.

The second product is the generalized matrix product or p-product [Rit91], where $p = t$:

$$[A_1(1), \ldots, A_g(1) | A_1(2), \ldots, A_g(2) | \ldots | A_1(g), \ldots, A_g(g)] \cdot_2 \begin{bmatrix} M_1(1) \\ \vdots \\ M_t(1) \\ \vdots \\ M_1(g) \\ \vdots \\ M_t(g) \end{bmatrix} :=$$

$$= \begin{bmatrix} A_1(1)M_1(1)+\ldots+A_g(1)M_1(g) & |A_1(2)M_1(1)+\ldots+A_g(2)M_1(g) & |\ldots & |A_1(g)M_1(1)+\ldots+A_g(2)M_1(g) \\ A_1(1)M_2(1)+\ldots+A_g(1)M_2(g) & |A_1(2)M_2(1)+\ldots+A_g(2)M_2(g) & |\ldots & |A_1(g)M_2(1)+\ldots+A_g(2)M_2(g) \\ \vdots & \vdots & & \vdots \\ A_1(1)M_t(1)+\ldots+A_g(1)M_t(g) & |A_1(2)M_t(1)+\ldots+A_g(2)M_t(g) & |\ldots & |A_1(g)M_t(1)+\ldots+A_g(2)M_t(g) \end{bmatrix}$$

This definition is extended by $\begin{bmatrix} \underline{A_1} \\ \underline{A_2} \end{bmatrix} \cdot_2 \underline{M} := \begin{bmatrix} \underline{A_1} \cdot_2 \underline{M} \\ \underline{A_2} \cdot_2 \underline{M} \end{bmatrix}$, etc.

The third product is a kind of scalar product in its first definition:

$$\begin{bmatrix} B_1(1) & B_1(2) & \ldots & B_1(g) \\ B_2(1) & B_2(2) & \ldots & B_2(g) \\ \vdots & \vdots & & \vdots \\ B_t(1) & B_t(2) & \ldots & B_t(g) \end{bmatrix} \cdot_3 \begin{bmatrix} y_1(1) & y_1(2) & \ldots & y_1(g) \\ y_2(1) & y_2(2) & \ldots & y_2(g) \\ \vdots & \vdots & & \vdots \\ y_t(1) & y_t(2) & \ldots & y_t(g) \end{bmatrix} :=$$

$$= \begin{bmatrix} B_1(1)y_1(1)+\ldots+B_1(g)y_1(g)+B_2(1)y_2(1)+\ldots+B_2(g)y_2(g)+\ldots+ \\ +B_t(1)y_t(1)+\ldots+B_t(g)y_t(g) \end{bmatrix}.$$

The extended definition is $\begin{bmatrix} \underline{A_1} \\ \underline{A_2} \end{bmatrix} \cdot_3 Y := \begin{bmatrix} \underline{A_1} \cdot_3 Y \\ \underline{A_2} \cdot_3 Y \end{bmatrix}$, etc.

LEMMA 5. *The two expressions, (17) and (18) are equal.*

PROOF: Use the definition above to get first

$$\underline{c} \cdot_2 \begin{bmatrix} M_1 \\ \vdots \\ M_t \end{bmatrix} \cdot_3 Y = [\underline{c}\cdot\underline{y}_1, \underline{c}\cdot\underline{y}_2, \ldots, \underline{c}\cdot\underline{y}_t] \begin{bmatrix} M_1 \\ \vdots \\ M_t \end{bmatrix},$$ second

$$[D_1,\ldots,D_g] \cdot_1 C \cdot_2 \underline{M} \cdot_3 Y = \sum_{k=1}^{g} D_k [\underline{c}(k)\underline{y}_1, \underline{c}(k)\underline{y}_2, \ldots, \underline{c}(k)\underline{y}_t] \begin{bmatrix} M_1(k) \\ \vdots \\ M_t(k) \end{bmatrix},$$

finally

$$D \cdot_1 C \cdot_2 \underline{M} \cdot_3 Y = \begin{bmatrix} \sum_{k=1}^{g} D_{1k}[\underline{c}(k)\underline{y}_1, \underline{c}(k)\underline{y}_2, \ldots, \underline{c}(k)\underline{y}_t] \begin{bmatrix} M_1(k) \\ \vdots \\ M_t(k) \end{bmatrix} \\ \vdots \\ \sum_{k=1}^{g} D_{gk}[\underline{c}(k)\underline{y}_1, \underline{c}(k)\underline{y}_2, \ldots, \underline{c}(k)\underline{y}_t] \begin{bmatrix} M_1(k) \\ \vdots \\ M_t(k) \end{bmatrix} \end{bmatrix}. \tag{19}$$

But the left hand side of (19) is equal to (17) by Lemma 4. ∎

§5.8. Dimension Two

In this section we give a detailed treatment to the case when A is a 2×2 matrix.

Let

$$\begin{bmatrix} x_1(n) \\ x_2(n) \\ 1 \end{bmatrix} := \begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(n-1) \\ x_2(n-1) \\ 1 \end{bmatrix}$$

In this case $\underline{\lambda} = [1, \lambda]$, where λ satisfies $(a_{11}+\lambda a_{21})\lambda = (a_{12}+\lambda a_{22})$. Let us introduce $\Delta := (a_{22} - a_{11})^2 + 4a_{12}a_{21}$. Therefore $\lambda_{1,2} = \frac{a22 - a11 \pm \sqrt{\Delta}}{2a_{21}}$ and $\mu = a_{11} + \lambda a_{21}$.

THEOREM 3. *The result of the n^{th} iteration of the affine transformation, $\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}$ is equal to*

$$\frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_2 & -\lambda_1 \\ -1 & 1 \end{bmatrix} \cdot \left(\begin{bmatrix} \mu_1^n & 0 \\ 0 & \mu_2^n \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} (\mu_1)_n & 0 \\ 0 & (\mu_2)_n \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right), \tag{20}$$

if $\Delta \neq 0$, and

$$\begin{bmatrix} \mu^n - \lambda a_{21} n \mu^{n-1} & -\lambda^2 a_{21} n \mu^{n-1} \\ a_{21} n \mu^{n-1} & \mu^n + \lambda a_{21} n \mu^{n-1} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} +$$

$$\begin{bmatrix} (\mu)_n - \lambda a_{21} (\mu)'_n & -\lambda^2 a_{21} (\mu)'_n \\ a_{21} (\mu)'_n & (\mu)_n + \lambda a_{21} (\mu)'_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \tag{21}$$

if $\Delta = 0$.

Recall, $(\mu)'_n = \mu^{n-1} + \mu^{n-2} + \ldots + \mu^2 + \mu + 1$, so $(\mu)'_n = (n-1)\mu^{n-2} + (n-2)\mu^{n-3} + \ldots + 2\mu + 1$. Of course, $(\mu^n)' = n\mu^{n-1}$.

PROOF: If $\Delta \neq 0$ (this is the non-defective case) we can use Theorem 2 to get $\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} =$

$$= \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix}^{-1} \left(\begin{bmatrix} \mu_1^n & 0 \\ 0 & \mu_2^n \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} (\mu_1)_n & 0 \\ 0 & (\mu_2)_n \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right),$$

or

$$\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_2 & -\lambda_1 \\ -1 & 1 \end{bmatrix} \cdot_1 \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} \cdot_2 \begin{bmatrix} (\mu_1)^n \\ (\mu_1)_n \\ (\mu_2)^n \\ (\mu_2)_n \end{bmatrix} \cdot_3 \begin{bmatrix} x_1(0) & x_2(0) \\ b_1 & b_2 \end{bmatrix}. \quad (20')$$

If $\Delta = 0$ (this corresponds to the defective case), we can use Theorem 1. We have $A(1) = [a_{22} - a_{21}\lambda]$, $\lambda(0) = [1, \lambda]$, $\lambda(1) = [1]$, $\mu(1) = a_{22} - a_{21}\lambda = a_{11} + a_{21}\lambda = \mu(0) =: \mu$,

$$R_n(0) = \mu^n [1, \lambda] \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + (\mu)_n [1, \lambda] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

(11) takes the form

$$x_1(n) = -\lambda x_2(n) + R_n(0),$$

$[1 \searrow 1]f = [\mu(1)]f = f(\mu)$ and

$$[1 \searrow 0]f = \underline{\lambda}(1) \cdot \underline{a}_{>1,1}(0) \cdot [\mu(1), \mu(0)]f = 1 \cdot a_{21} \cdot [\mu, \mu]f =^{(by (6))} a_{21} f'(\mu).$$

Since $f(\mu)$ is either μ^n or $(\mu)_n$, we have by Theorem 1. that

$$x_2(n) = \mu^n x_2(0) + (\mu)_n b_2 + a_{21} n \mu^{n-1} [1, \lambda] \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + a_{21} (\mu)'_n [1, \lambda] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} =$$

$$= \left[a_{21} n \mu^{n-1}, \ \lambda a_{21} n \mu^{n-1} + \mu^n \right] \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \left[a_{21} (\mu)'_n, \ \lambda a_{21} (\mu)'_n + (\mu)_n \right] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Since the coefficient of $x_2(0)$ or b_2 for $x_1(n)$ is
$-\lambda\left(\lambda a_{21}f'(\mu)+f(\mu)\right)+\lambda f(\mu)=-\lambda^2 a_{21}f'(\mu)$, we obtain (21) or

$$\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} = \begin{bmatrix} 1 & -\lambda & 0 & -\lambda^2 \\ 0 & 1 & 1 & \lambda \end{bmatrix} \cdot_2 \begin{bmatrix} (\mu)^n \\ (\mu)_n \\ a_{21} n \mu^{n-1} \\ a_{21}(\mu)'_n \end{bmatrix} \cdot_3 \begin{bmatrix} x_1(0) & x_2(0) \\ b_1 & b_2 \end{bmatrix}. \tag{21'}$$

∎

Specialization of these formulas leads to the usual expressions for recurrence sequences in number theory. Let us see how:

$$\begin{bmatrix} x(n) \\ x(n-1) \\ 1 \end{bmatrix} := \begin{bmatrix} a_{11} & a_{12} & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(n-1) \\ x(n-2) \\ 1 \end{bmatrix}$$

In this special case $\lambda_1=-\mu_2$ and $\lambda_2=-\mu_1$.

If $\Delta \neq 0$ then

$$x(n)=\frac{(\mu_1)^n-(\mu_2)^n}{\mu_1-\mu_2}\cdot x(1)-\mu_1\mu_2\frac{(\mu_1)^{n-1}-(\mu_2)^{n-1}}{\mu_1-\mu_2}\cdot x(0)+\sum_{j=1}^{n-1}\frac{(\mu_1)^j-(\mu_2)^j}{\mu_1-\mu_2}\cdot b_1. \tag{22}$$

On the other hand, if $\Delta = 0$

$$x(n)=n\mu^{n-1}\cdot x(1)-(n-1)\mu^n\cdot x(0)+((n-1)\mu^{n-2}+\ldots+2\mu+1)\cdot b_1. \tag{23}$$

Finally, if we take $b_1=0$ and replace $x(1)$ and $x(0)$ by α_1 and α_2 satisfying

$$\begin{cases} 1\cdot\alpha_1+1\cdot\alpha_2=x(0) \\ \mu_1\cdot\alpha_1+\mu_2\cdot\alpha_2=x(1), \end{cases} \quad \textit{when } \Delta\neq 0,$$

$$\begin{cases} 1\cdot\alpha_1+0\cdot\alpha_2=x(0) \\ \mu\cdot\alpha_1+\mu\cdot\alpha_2=x(1), \end{cases} \quad \textit{when } \Delta = 0,$$

we reach the formula found, for instance, in [Niv91].

$$x(n) = \begin{cases} \alpha_1 \cdot \mu_1^n + \alpha_2 \cdot \mu_2^n & \text{if } \Delta \neq 0 \\ (\alpha_1 + n\alpha_2)\mu^n & \text{if } \Delta = 0 \end{cases}$$

The Fibonacci sequence corresponds to the initial condition $[x(1), x(0)] = [1, 0]$, while the Lucas numbers come from setting $[x(1), x(0)] = [a_{11}, 2]$.

CHAPTER 6

DETERMINANTS, POWER SERIES, PARTITIONS

§6.1. Introduction, Notation

The purpose of this chapter is threefold. First we unify the handling of convolutions of sequences. This naturally leads to determinants of *lower Hessenberg* matrices, i.e. matrices whose entries above the superdiagonal are all zeros. Next, in section 3, we show how the major combinatorial numbers, for example the binomial coefficients, the Stirling numbers of either kind and the Lah numbers can be assembled in special Hessenberg determinants.

Our final aim is to show how easily we can manipulate formal power series using these determinants and by doing so, we also hope to rescue from oblivion some beautiful formulas of the last century. These *Hessenberg series representations* [our term] are not of theoretical interest only: a Hessenberg matrix is sparse and its determinant can be evaluated relatively cheaply [see the beginning of section 4]. This section can be read almost independently from the rest of the paper, so the reader who would like to introduce interesting formulas in his or her calculus class can jump forward.

We will use the following notation. Given n numbers $\{x_j\}_{j=1}^{n}$, let

$$G(t) := (1 - tx_1) \cdot \ldots \cdot (1 - tx_n),$$

$$F(t) := c_0(t - x_1) \cdot \ldots \cdot (t - x_n) = c_0 t^n + c_1 t^{n-1} + \ldots + c_{n-1}t + c_n.$$

The elementary, complete homogeneous and power sum symmetric functions over $x_1, \dots, x_n$ are defined by

$$\sigma_r := \sum_{j_1 < \dots < j_r} x_{j_1} \cdots x_{j_r} = (-1)^r G(t)//t^r = (-1)^r \frac{c_r}{c_0},$$

$$h_r := \sum_{j_1 \le \dots \le j_r} x_{j_1} \cdots x_{j_r} = \sum_{\sum k_j = r} x_1^{k_1} \cdots x_n^{k_n} = \frac{1}{G(t)}//t^r,$$

$$p_r := x_1^r + \dots + x_n^r = -r \cdot \log G(t)//t^r.$$

We prefer $P//t^r$ to denote the coefficient of t^r in the formal power series P to the usual $[t^r]P$, since $P \cdot t^k//t^{k+r} = P//t^r$ holds. Let us agree that h_0, p_0, σ_0 are all 1.

We ask a little more patience from the reader, since we have to make some important definitions before we can start our discussion.

For the unsigned Stirling numbers of the first kind and for the Stirling numbers of the second kind we will use $\begin{bmatrix} n \\ r \end{bmatrix}$ and $\begin{Bmatrix} r \\ n \end{Bmatrix}$ respectively. This notation was introduced by D. Knuth [Knu92]. We define them as special cases with $q = 1$ of

$$\begin{bmatrix} n \\ r \end{bmatrix}_q := \sigma_{n-r}\left(x_1 := 0, x_j := 1 + q + \dots + q^{j-2}\right) \quad \text{for the first (cycle) kind,} \quad (1)$$

$$\begin{Bmatrix} r \\ n \end{Bmatrix}_q := h_{r-n}\left(x_j := 1 + q + \dots + q^{j-1}\right) \quad \text{for the second (subset) kind.} \quad (2)$$

Although these definitions look unusual, if $q = 1$, by a change of parameters, we obtain

$$\begin{bmatrix} n+1 \\ n+1-r \end{bmatrix} = \sigma_r(x_j := j) \text{ and } \begin{Bmatrix} n+r \\ n \end{Bmatrix} = h_r(x_j := j), \quad (3)$$

which is [Knu92, (2.6)] (or see [Com74, pp. 214,207]).

Remark that H. Gould [Gou61] originally introduced the q-Stirling numbers $S_1(n,r,q)$ and $S_2(n,r,q)$ by (3), therefore his definitions differ from ours.

It is instructive to compare definitions (1) and (2) with expressions for the q-binomial coefficients (also called *Gaussian polynomials*):

$$\sigma_{n-r}\left(x_1 := 0, x_j := q^{j-2}\right) = q^{\binom{n-r}{2}}\binom{n-1}{r-1}_q \text{ and } h_{r-n}\left(x_j := q^{j-1}\right) = \binom{r-1}{n-1}_q,$$

where

$$\binom{m}{\ell}_q := \frac{(1-q^m)(1-q^{m-1})\cdots(1-q^{m-\ell+1})}{(1-q)(1-q^2)\cdots(1-q^\ell)}.$$

[Just divide both the numerator and the denominator by $(1-q)^\ell$ to see that $\binom{m}{\ell}_{q\to 1} = \binom{m}{\ell}$.]

TABLE 6-1. VARIOUS SUBSTITUTIONS FOR x_j

	p_r	σ_r	h_r
$\{x_j := 1\}_{j=1}^n$	n	$\binom{n}{n-r}$	$\binom{n+r-1}{n-1}$
$\{x_j := q^{j-1}\}_{j=1}^n$	$\frac{1-q^{rn}}{1-q^r}$	$\binom{n}{n-r}_q$	$\binom{n+r-1}{n-1}_q$
$\{x_j := q^{j-1}\}_{j=1}^\infty$	$\frac{1}{1-q^r}$	$\frac{1}{(1-q)\cdots(1-q^r)}$	$\frac{q^{\binom{r}{2}}}{(1-q)\cdots(1-q^r)}$
$\{x_j := j-1\}_{j=1}^n$		$\left[{n \atop n-r}\right]$	$\left\{{n+r-1 \atop n-1}\right\}$
$\{x_j := \frac{1-q^{j-1}}{1-q}\}_{j=1}^n$		$\left[{n \atop n-r}\right]_q$	$\left\{{n+r-1 \atop n-1}\right\}_q$

§6.2. Convolution of Sequences

Given two sequences $\{a_r\}$ and $\{b_r\}$, finite or infinite, we frequently are able to find coefficients $g_{r,j}$ such that

$$\sum_{s=0}^{r} g_{r,s}a_{r-s}b_s = g_{r,0}a_rb_0 + g_{r,1}a_{r-1}b_1 + \ldots + g_{r,r-1}a_1b_{r-1} + g_{r,r}a_0b_r = K_r. \quad (4)$$

To convince ourselves of the general nature of (4) we collected a few examples in the following table.

TABLE 6-2. EXAMPLES OF CONVOLUTIONAL IDENTITIES

	a_s	b_s	$g_{r,s}$	K_r
(5)	c_s	p_s	$r\delta_{s0} + 1 \cdot (1 - \delta_{s0})$	0
(6)	c_s	h_s	1	0
(7)	h_s	p_s	$r\delta_{s0} - 1 \cdot (1 - \delta_{s0})$	0
(8)	1	$B_s(x)$	$\binom{r+1}{s}$	$(r+1)x^r$
(9)	$f(x)^{(s)}$	$g(x)^{(s)}$	$\binom{r}{s}$	$(f(x) \cdot g(x))^{(r)}$
(9')	$P_s(y)$	$P_s(x)$	$\binom{r}{s}$	$P_r(x+y)$
(10)	C_s	C_s	1	C_{r+1}
(10')	$P_s(y)$	$P_s(x)$	1	$\frac{P_{r+1}(x)-P_{r+1}(y)}{x-y}$
(10")	$P_{r-s}(y)$	$P_s(x)$	1	$K_r \frac{P_{r+1}(x)P_r(y)-P_{r+1}(y)P_r(x)}{x-y}$

Identities (5),(6),(7) are true for $1 \leq r \leq n$ and they belong to *Newton, Wroński* and *Brioschi* respectively.

The relationship (8) for the *Bernoulli* polynomials $B_j(x)$ is valid for $r > 0$ ($B_0 := 1$). If we write the indices as exponents, (8) can be abbreviated as $(B(x) + 1)^{r+1} - B(x)^{r+1} = (x^{r+1})'$. The Bernoulli numbers can be obtained by substituting $x := 0$. They satisfy

$$p_r(x_j := j-1) = \frac{\sum_{j=0}^{r} \binom{r+1}{j} B_j n^{r+1-j}}{r+1} = \text{(formally)}\ \frac{(B+n)^{r+1} - B^{r+1}}{r+1}. \quad (11)$$

The *Faulhaber numbers* $A_j^{(r)}$, introduced by D. Knuth recently [Knu93] give a further example of (4). They are related to (8) and connected by the equation $\sum_{j=0}^{k} \binom{r-j}{2k+1-2j} A_j^{(r)} = 0$, where $k > 0$ and $A_0^{(r)} = 1$.

Formula (9) is, of course Leibnitz's. Identity (10) is satisfied by the *Catalan numbers,* where $C_0 := 1$ and

$$C_r := \frac{1}{r+1}\binom{2r}{r} = \frac{1-\sqrt{1-4x}}{2x} // x^r.$$

Polynomials that satisfy (9') or (10) are called of binomial type and Newtonian respectively. The former ones are related to umbral calculus [Rom78], while those satisfying (10') lead to divided differences, Vandermonde determinants [Hir92]. Equation (10") is the *Christoffel-Darboux formula* for orthonormal polynomials.

In our last example we would like to rewrite a sum as a product, or the way around [And76, page 98]:

$$L(q) := \sum_{n=0}^{\infty} a_n q^n = \frac{1}{\prod_{n=1}^{\infty}(1-q^n)^{b_n}}.$$

On one hand,

$$q \cdot \frac{L'(q)}{L(q)} = \frac{\sum n a_n q^n}{\sum a_n q^n},$$

on the other hand,

$$q \cdot [\log L(q)]' = q \sum_{n=1}^{\infty} \frac{n b_n q^{n-1}}{1-q^n} = \sum_{n=1}^{\infty} n b_n \left(\sum_{m=1}^{\infty} q^{mn} \right) =$$

$$\sum_{N=1}^{\infty} \left(\sum_{n|N} n b_n \right) q^N =: \sum_{N=1}^{\infty} d_N q^N.$$

Therefore sequences $\{d_r\}$ and $\{a_r\}$ are connected by an equation of type (7), namely

$$N a_N - d_1 a_{N-1} - \ldots - d_{N-1} a_1 - d_N a_0 = 0.$$

After the examples, it is high time we solved (4) for b_r (or a_r). Write

$$\begin{aligned}
&g_{1,1} a_0 b_1 &&= -g_{1,0} a_1 b_0 + K_1,\\
&g_{2,1} a_1 b_1 + g_{2,2} a_0 b_2 &&= -g_{2,0} a_2 b_0 + K_2,\\
&\quad\vdots \\
&g_{r,1} a_{r-1} b_1 + g_{r,2} a_{r-2} b_2 + \ldots \quad + g_{r,r} a_0 b_r &&= -g_{r,0} a_r b_0 + K_r.
\end{aligned}$$

To obtain b_r, use Cramér's rule, then bring the last column forward; a_r can be obtained similarly.

$$b_r = \frac{\begin{vmatrix} g_{1,0}b_0a_1 - K_1 & g_{1,1}a_0 & & \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & & g_{r-1,r-1}a_0 \\ g_{r,0}b_0a_r - K_r & g_{r,1}a_{r-1} & \cdots & g_{r,r-1}a_1 \end{vmatrix}}{(-a_0)^r \cdot g_{1,1}g_{2,2}\cdots g_{r,r}}. \tag{12}$$

The numerators in (12) are determinants of *lower Hessenberg* matrices. Let us remark that every real matrix can be transformed into a Hessenberg matrix using a succession of *Householder* transformations (*Weyl* reflections).

Newton's formula, i.e. (5) and the identity

$$F(t) = c_0t^n + c_1t^{n-1} + \ldots + c_n = t^n \cdot \left[c_n\left(\frac{1}{t}\right)^n + \ldots + c_1\left(\frac{1}{t}\right) + c_0\right]$$

give us for the roots $\{x_1, \ldots, x_n\}$ of $F(t)$

$$p_r = x_1^r + \ldots + x_n^r = \frac{1}{(-c_0)^r} \cdot \begin{vmatrix} c_1 & c_0 & & & \\ 2c_2 & \ddots & \ddots & & \\ 3c_3 & c_2 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & c_0 \\ rc_r & c_{r-1} & \cdots & c_2 & c_1 \end{vmatrix}, \tag{13}$$

and

$$\frac{1}{x_1^r} + \ldots + \frac{1}{x_n^r} = \frac{1}{(-c_n)^r} \cdot \begin{vmatrix} c_{n-1} & c_n & & & \\ 2c_{n-2} & \ddots & \ddots & & \\ 3c_{n-3} & c_{n-2} & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & c_n \\ rc_{n-r} & c_{n-r+1} & \cdots & c_{n-2} & c_{n-1} \end{vmatrix}, \tag{13'}$$

provided c_0 and c_n are not zero. Using (13') we can express, for example, $\frac{1}{1^r}+\frac{1}{2^r}+\ldots+\frac{1}{n^r}$ by the Stirling numbers of the first kind.

Several determinantal expressions exist for the Bernoulli numbers. If we apply (12), we can obtain a possibly new one from (8):

$$B_r = (-1)^r \frac{\begin{vmatrix} \binom{2}{0} & \binom{2}{1} & & & \\ \vdots & & & \ddots & \\ \vdots & & & & \binom{r}{r-1} \\ \binom{r+1}{0} & \binom{r+1}{1} & \cdots & \cdots & \binom{r+1}{r-1} \end{vmatrix}}{(r+1)!}. \tag{14}$$

Let us mention that (11) can be written as

$$p_r(x_j := j-1) = (-1)^r \frac{\begin{vmatrix} n & \binom{1}{0} & & & \\ n^2 & \binom{2}{0} & \binom{2}{1} & & \\ \vdots & & & \ddots & \\ \vdots & & & & \binom{r}{r-1} \\ n^{r+1} & \binom{r+1}{0} & \binom{r+1}{1} & \cdots & \binom{r+1}{r-1} \end{vmatrix}}{(r+1)!}. \tag{11'}$$

By exchanging the rôle of $\{a_r\}$ and $\{b_r\}$ in (5), we can get

$$\sigma_r = \frac{1}{r!} \cdot \begin{vmatrix} p_1 & 1 & & & \\ p_2 & \ddots & 2 & & \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \text{r-1} \\ p_r & \cdots & \cdots & p_2 & p_1 \end{vmatrix}. \tag{15}$$

If A is a $n \times n$ matrix, its characteristic polynomial $\det(tI - A)$ is equal to $t^n + \sum_{r=1}^{n}(-1)^r \sigma_r t^{n-r}$ by the Cayley-Hamilton theorem. For example, $\sigma_1 = tr(A) = p_1$

and $\sigma_n = \det(A)$. Therefore we can apply (15) to evaluate it. Naturally, here $\{x_j\}, j = 1, \ldots, n$ are the eigenvalues of A, and $p_r = tr(A^r) = \sum_{j=1}^{n} x_j^r$.

Identity (15) and similar ones obtained from (5),(6),(7) are not new, see [Mac79, page 20] or [Kri86, page 55]. When we apply (12) to these three identities, we have $K_j = 0$ and $b_0 = 1$. We will change the meaning of c_r from now on, namely, set $c_r := g_{r,r}a_0$. Let

$$det_m[g_{rs}a_{r-s}] := \begin{vmatrix} g_{1,0}a_1 & c_1 & & \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & & c_{m-1} \\ g_{m,0}a_m & g_{m,1}a_{m-1} & \cdots & g_{m,m-1}a_1 \end{vmatrix}. \tag{16}$$

Note that we index the rows r of the $m \times m$ matrix from 1 to m, while its columns s from 0 to $m-1$. If $m \le n$, we have

TABLE 6-3. DETERMINANTAL RELATIONSHIPS

c_r	$g_{r,s}$	a_j, if $j := r - s$	$\det_m$
$-r$	1	p_j	$m! \cdot h_m$
r	1	p_j	$m! \cdot \sigma_m$
1	1	h_j	σ_m
1	1	σ_j	h_m
1	$1+(r-1)\delta_{s0}$	σ_j	p_m
1	$1+(r-1)\delta_{s0}$	h_j	$(-1)^{m-1}p_m$

From Table 1. and 3. we can gain thirty identites; the expression for $p_r(x_j := j-1)$ will be different from (11'). In the next section we will build a three level hierarchy among the terms in the expansion of det_m, the third level will be the determinant itself.

§6.3. Hessenberg Determinants and Combinatorial Numbers

We will expand the determinant (16) and reorder its 2^{m-1} terms. As an illustration, take the $m = 4$ case:

$$\begin{vmatrix} g_{10}a_1 & c_1 & & \\ g_{20}a_2 & g_{21}a_1 & c_2 & \\ g_{30}a_3 & g_{31}a_2 & g_{32}a_1 & c_3 \\ g_{40}a_4 & g_{41}a_3 & g_{42}a_2 & g_{43}a_1 \end{vmatrix}$$

Its expansion is

$$\begin{array}{cccccccc} - & c_1c_2c_3 & & g_{40} & & a_4 \\ + & c_1c_2 & g_{30} & & g_{43} & a_3a_1 \\ + & c_2c_3 & g_{10} & & g_{41} & a_3a_1 \\ + & c_1c_3 & g_{20} & & g_{42} & a_2^2 \\ - & c_1 & g_{20} & g_{32} & g_{43} & a_2a_1^2 \\ - & c_2 & g_{10} & g_{31} & g_{43} & a_2a_1^2 \\ - & c_3 & g_{10} & g_{21} & g_{42} & a_2a_1^2 \\ + & & g_{10} & g_{21}g_{32} & g_{43} & a_1^4 \end{array}$$

The coefficient of $(a_1^{k_1} \cdots a_m^{k_m})$ in a general term (where $k_j \geq 0$) is of the form $t[c_r; g_{r,s}] := (-1)^{k_0} \cdot (k_0 \text{ of the } c\text{'s }) \cdot g_{j,0} \cdot (\ell - 2 \text{ of the } g\text{'s }) \cdot g_{m,i}$, where

$$\ell := k_1 + \ldots + k_m \tag{17}$$

and $k_0 := m - \ell \geq 0$. Special cases of the following lemma play an important rôle in calculations:

Lemma 1. *Let h, h_1, h_2 be arbitrary functions from $\{0, 1, \ldots, m\}$ to $\mathbb{C} - \{0\}$. We have*

$$t\left[h(0)c_r; h(r-s)g_{r,s}\right] = h(0)^{k_0}h(1)^{k_1} \cdots h(m)^{k_m} \cdot t[c_r; g_{r,s}], \tag{a}$$

and

$$t\left[\frac{h_2(r)}{h_1(r)}c_r; g_{r,s}\right] = \frac{h_2(0)h_2(1)\cdots h_2(m-1)}{h_1(1)h_1(2)\cdots h_1(m)} \cdot t\left[c_r; \frac{h_1(r)}{h_2(s)}g_{r,s}\right]. \tag{b}$$

In particular, we have

$$t[c_r; (r-s)g_{rs}] = 1^{k_1}2^{k_2}\cdots m^{k_m} \cdot t[c_r; g_{rs}], \tag{a'}$$

$$t[c_r; (r-s-1)!g_{rs}] = (0!)^{k_1}(1!)^{k_2}\cdots(m-1)^{k_m} \cdot t[c_r; g_{rs}], \tag{a''}$$

and

$$t\left[c_r; \frac{(r-1)!}{s!}g_{rs}\right] = t[rc_r; g_{rs}] \quad \blacksquare \tag{b'}$$

The following lemmas will have their applications in the next section.

LEMMA 2.

$$\frac{1}{m!}det_m\left[\binom{r}{s}(r-s)!g_{rs}a_{r-s}\right] = det_m\left[g_{rs}a_{r-s}\right].$$

PROOF: Since $t\left[c_r; \frac{r!}{s!}g_{rs}a_{r-s}\right] = t[rc_r; rg_{rs}a_{r-s}]$ by (b') of the previous lemma, we can factor out r from row r, $r = 1, \ldots, m$.

LEMMA 3. *We can change the dimension of the matrix by the formula*

$$det_{m+1}[b_{r-1}\delta_{s0} + a_{r-s}(1-\delta_{s0})] = det_m[(b_0a_r - b_ra_0)\delta_{s0} + a_{r-s}(1-\delta_{s0})] \quad \blacksquare$$

LEMMA 4.

$$\frac{1}{m!}det_{m+1}\left[b_{r-1}\delta_{s0} + \binom{r-1}{s-1}(r-s)!a_{r-s}(1-\delta_{s0})\right] = det_{m+1}\left[b_{r-1}\delta_{s0} + a_{r-s}(1-\delta_{s0})\right].$$

PROOF: $LHS = \frac{1}{m!}det_m\left[\binom{r}{s}(r-s)!\left((b_0a_r - b_ra_0)\delta_{s0} + a_{r-s}(1-\delta_{s0})\right)\right]$, $RHS = det_m[(b_0a_r - b_ra_0)\delta_{s0} + a_{r-s}(1-\delta_{s0})]$ by Lemma 3., and they are equal by Lemma 2. $\blacksquare$

Let us call those terms which have the same $(a_1^{k_1}\cdots a_m^{k_m})$ factors a *cluster* (or type). Since

$$\sum_{j=1}^{m} jk_j = m, \tag{18}$$

the elements of a cluster belong to the same partition of the integer m. Next regard the *superclusters* of terms for which ℓ is the same. In the $m = 4$ illustration above the two a_3a_1 and the a_2^2 term belong to the same $\ell = 2$ supercluster.

By carefully choosing the coefficients $c_r, f_r, g_{r,s}$ and calculating the cluster sums, we obtain numbers which are well known in combinatorics.

TABLE 6-4. SPECIAL HESSENBERG DETERMINANTS

	c_r	$g_{r,s}$	$S_1 = \sum_{(k_1,\ldots,k_m)=fixed}$	$S_2 = \sum_{\ell=const}$	$S_3 = \sum_{\ell=1}^{m}$
(19)	-1	1	$\frac{(k_1+\ldots+k_m)!}{k_1!\ldots k_m!}$	$\binom{m-1}{\ell-1}$	2^{m-1}
(20)	-1	$r-s$	$\frac{(k_1+\ldots+k_m)!}{k_1!\ldots k_m!}\cdot 1^{k_1}\ldots m^{k_m}$	$\binom{m+\ell-1}{2\ell-1}$	$\frac{t}{1-3t+t^2}//t^m$
(21)	-1	$1+(r-1)\cdot\delta_{s0}$	$m\cdot\frac{(k_1+\ldots+k_m-1)!}{k_1!\ldots k_m!}$	$\binom{m}{\ell}$	2^m-1
(22)	-1	$\binom{r-1}{s}$	$\frac{m!}{k_1!\ldots k_m!\cdot(1!)^{k_1}\ldots(m!)^{k_m}}$	$bell_{m,\ell}$	Y_m
(22')	-1	$\binom{r-1}{s}$	$\frac{m!}{k_1!\ldots k_m!\cdot(1!)^{k_1}\ldots(m!)^{k_m}}$	$\left\{{m \atop \ell}\right\}$	$Bell_m$
(23)	$-r$	1	$\frac{m!}{k_1!\ldots k_m!\cdot 1^{k_1}\ldots m^{k_m}}$	$\left[{m \atop \ell}\right]$	$m!$
(24)	-1	$(r-s)\binom{r-1}{s}$	$\frac{m!}{k_1!\ldots k_m!\cdot(0!)^{k_1}\ldots((m-1)!)^{k_m}}$	$\ell^{m-\ell}\binom{m}{\ell}$	
(25)	$-r$	$r-s$	$\frac{m!}{k_1!\ldots k_m!}$	$\frac{m!}{\ell!}\binom{m-1}{\ell-1}$	

The numbers in column S_1 are the sums of coefficients belonging to $(a_1^{k_1}\cdots a_m^{k_m})$. To obtain the numbers in column S_2, first take $a_1 := a_2 := \ldots a_n := 1$ (*except* in (22)), sum for those m-tuplets $(k_1,\ldots,k_m)$ which satisfy (17) and (18), and finally apply known identities, which can be found in [Ego77, p.184], for example.

Let us give an example for the usage of Table 4. The number of permutations [partitions] of a set of m elements that have k_j cycles [classes] of length [cardinality]

j is the quantity in (23,S1) [(22',S1)] respectively, and if we sum these numbers for those m-tuplets $(k_1, \ldots, k_m)$ which satisfy (17) and (18) we obtain $\left[{m \atop \ell}\right]$ and $\left\{{m \atop \ell}\right\}$.

Instead of giving a rigorous proof of the table here, we just indicate that once we establish (19),(21) and (22'), the rest can be obtained using the mentioned special cases of Lemma 1.

If we do *not* substitute $a_j := 1$ in (22), we obtain the *partial* Bell polynomials $bell_{m,\ell}$ in S_2. Their sum in S_3 is the *complete* Bell polynomial Y_m. (The Bell *numbers* were marked as $Bell_m$ in Table 4.) This way we proved the fact mentioned in [Com74, p.135] that

$bell_{m,\ell}(a_j := 1) = \left\{{m \atop \ell}\right\}$, $bell_{m,\ell}(a_j := j) = \ell^{m-\ell}\binom{m}{\ell}$ (idempotent number).
$bell_{m,\ell}(a_j := (j-1)!) = \left[{m \atop \ell}\right]$, $bell_{m,\ell}(a_j := j!) = \frac{m!}{\ell!}\binom{m-1}{\ell-1}$ (Lah number),

We have two remarks to make at this point. First, if we want to assemble $\left[{m \atop \ell}\right]_q$, then (by definition (1)), it is enough to change $c_r := r$ to $c_r := 1 + q + \ldots + q^{r-1}$ in (23). Second, since Hessenberg determinantal representations are not necessarily unique, we can set our additional goal to find the most aesthetic one.

We can use the table to obtain identities like

$$\sum_{\sum k_j = \sum j k_j = m} \frac{(k_1 + \ldots + k_m)!}{k_1! \ldots k_m!} (-a_0)^{k_0} a_1^{k_1} \cdots a_m^{k_m}$$

$$= \det \begin{bmatrix} a_1 & a_0 & & & \\ a_2 & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ a_{m-1} & & \ddots & \ddots & a_0 \\ a_m & a_{m-1} & \ldots & a_2 & a_1 \end{bmatrix}, \quad \text{(26, from (19))}$$

$$\sum_{\sum k_j = \sum j k_j = m} m \cdot \frac{(k_1 + \ldots + k_m - 1)!}{k_1! \ldots k_m!} (-a_0)^{k_0} a_1^{k_1} \cdots a_m^{k_m}$$

$$= \det \begin{bmatrix} a_1 & a_0 & & & \\ 2a_2 & \ddots & \ddots & & \\ \vdots & a_2 & \ddots & \ddots & \\ (m-1)a_{m-1} & & \ddots & \ddots & a_0 \\ m a_m & a_{m-1} & \cdots & a_2 & a_1 \end{bmatrix} \qquad \text{(27, from (21))}$$

$$\sum_{\sum k_j = \sum j k_j = m} \frac{m!}{k_1! \ldots k_m!} \left(\frac{f_1}{1}\right)^{k_1} \cdots \left(\frac{f_m}{m}\right)^{k_m} (-a_0)^{k_0} a_1^{k_1} \cdots a_m^{k_m}$$

$$= \det \begin{bmatrix} f_1 a_1 & a_0 & & & \\ f_2 a_2 & \ddots & 2a_0 & & \\ \vdots & f_2 a_2 & \ddots & \ddots & \\ f_{m-1} a_{m-1} & & \ddots & \ddots & (m-1)a_0 \\ f_m a_m & f_{m-1} a_{m-1} & \cdots & f_2 a_2 & f_1 a_1 \end{bmatrix} . \qquad \text{(28, from (23))}$$

Formulas (26), (27) and (28), (with $a_0 = \pm 1$, $f_j = 1$ or j) can be found in [Mui28, pp. 703,712] (with some misprints). Let us remark two things. First, if m is a prime number, expression (27) $\equiv a_1^m \equiv a_1 \pmod{m}$. Second, if we take $a_0 = -1, a_1 = \ldots = a_n = 1$ in (28), we get $m!\cdot$(cycle index of the symmetric group S_m) (for a definition, see [Pol37).

§6.4. Operations with Power Series Using Hessenberg Determinants

Expression (12) was obtained by Cramér's rule, therefore it can be looked upon with great suspicion by computation-conscious people. But our matrices are sparse

and by expanding them by their last row, we see that they obey a rule of type (4) too, namely

$$\det M_k = g_{k,k-1}a_1 \det M_{k-1} - g_{k,k-2}c_{k-1}a_2 \det M_{k-2} + \ldots + \\ + (-1)^{k-2} g_{k,1}c_{k-1}\cdots c_2 a_{k-1} \det M_1 + (-1)^{k-1} g_{k,0}c_{k-1}\cdots c_2 c_1 a_k. \quad (29)$$

This makes their computation possible in $O(k^3)$ steps, a much less formidable number than the 2^{k-1} terms in their expansion would suggest. Moreover, if we possess $O\left(k^4\right)$ processors, they can be evaluated in $O\left((\log k)^2\right)$ steps [Lak90, page 443].

The calculus of residues [Ego77] is another (and powerful) tool to describe operations on series, but the Hessenberg series approach is easier to use, in our opinion.

Let us start our treatment of the power series with *Faà di Bruno*'s formula (1855) about the m^{th} power of composite functions. This formula survived in the modern literature [Mel73, page 214]. Look at (22) in the second table to obtain

$$(g \circ f)^{(m)}(x) = \\ \sum_{\ell=1}^{m} \left(\sum_{\sum k_j = \ell, \sum j k_j = m} \frac{m!}{k_1! \ldots k_m! \cdot (1!)^{k_1} \ldots (m!)^{k_m}} \left[f'(x)\right]^{k_1} \cdots \left[f^{(m)}(x)\right]^{k_m} \right) g^{(\ell)}(f(x)) = \\ = \det \begin{bmatrix} f'(x)D & -1 & & & \\ f''(x)D & \ddots & -1 & & \\ \vdots & & \ddots & \ddots & \\ \binom{m-2}{0} f^{(m-1)}(x)D & & \ddots & \ddots & -1 \\ \binom{m-1}{0} f^{(m)}(x)D & \binom{m-1}{1} f^{(m-1)}(x)D & \ldots & \binom{m-1}{m-2} f''(x)D & \binom{m-1}{m-1} f'(x)D \end{bmatrix} \\ \text{applied to } g(f(x)), \quad (30)$$

where D is the differentiation operator, i.e. $D^s f^{(k)}(x) := f^{(s+k)}(x)$.

We can use this formula to substitute a power series into another one. If $f(x)$ and $g(x)$ are power series, namely $f(x) = \sum_{k=1}^{\infty} a_k x^k$ and $g(x) = \sum_{k=0}^{\infty} b_k x^k$ we get

$$h(x) := (g \circ f)(x) = b_0 + \sum_{m=1}^{\infty} \frac{(g \circ f)^{(m)}(0)}{m!} x^m = b_0 + \sum_{m=1}^{\infty} c_m x^m,$$

where

$$c_m = \sum_{\sum jk_j = m} \frac{(k_1 + \ldots + k_m)!}{k_1! \cdots k_m!} a_1^{k_1} \cdots a_m^{k_m} b_{k_1+\ldots+k_m} =$$

$$= \det \begin{bmatrix} a_1 X & -1 & & & \\ a_2 X & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ a_{m-1} X & & \ddots & \ddots & -1 \\ a_m X & a_{m-1} X & \ldots & a_2 X & a_1 X \end{bmatrix} b, \quad (31)$$

where X is the indexing operator, $X^s b_k = b_{k+s}$.

Let us now discuss the (compositional) inverse of a power series. Let $w = f(z) := z \sum_{k=0}^{\infty} a_k z^k$ and let its inverse be $z = g(w) := w \sum_{k=0}^{\infty} b_k w^k$ (notice that we shifted the indexing). Then [Dav66, page 273]

$$b_m = (-1)^m \frac{1}{a_0^{2m+1}} \sum_{\sum k_j = \sum jk_j = m} \frac{(m + 1 + k_1 + \ldots + k_m - 1)!}{(m+1)! k_1! \cdots k_m!} (-a_0)^{k_0} a_1^{k_1} \cdots a_m^{k_m} \quad (32)$$

For example,

$$b_4 = \frac{1}{a_0^9} \left(-a_0^3 a_4 + 6a_0^2 a_1 a_3 + 3a_0^2 a_2^2 - 21 a_0 a_1^2 a_2 + 14 a_1^4 \right). \quad (33)$$

Observe that the coefficient of a_1^m is the Catalan number C_m. If we set $a_0 := -1$ and $a_1, \ldots, a_n := x$, the S_2 sum we obtain is an orthogonal polynomial $b_m(x)$, since it satisfies the three-term recurrence

$$(m+2) \cdot b_{m+1}(x) - (2m+1)(2x+1) \cdot b_m(x) + (m-1) b_{m-1}(x) = 0,$$

with the initial condition $b_0(x) := 1$ and $b_1(x) := x$. After a little calculation we notice that $b_m(x)$ is a Gegenbauer-type orthogonal polynomial, therefore we can obtain its generating function

$$b_m(x) = \frac{1}{\binom{m+1}{2}} \cdot \frac{x}{(1-2(2x+1)w+w^2)^{\frac{3}{2}}} // w^{m-1}.$$

We now rewrite the previous expression in a Hessenberg determinantal form. Actually, let us try to solve the following, more general right inverse problem : given the power series h and f, determine g so that $g \circ f \equiv h$.

Suppose $h(z) = z \sum_{m=0}^{\infty} c_m z^m$. (If $c_0 = 1$ and $c_j = 0$, $j = 1, 2, \ldots$, we have the original problem about the compositional inverse). We can use (31) to set up a system of linear equations, then solve it for b_m (like we did in the beginning of section 2). The denominator is $(-1)^m a_0^{\frac{(m+2)(m+1)}{2}}$; the entry of the numerator determinant at row r, $(r = 1, 2, \ldots, m+1)$, column $s = 0$ is c_{r-1}, while for $(s = 1, 2, \ldots, m)$ it is

$$\sum_{\sum jk_j = r-s, \sum k_j = s} \frac{s!}{k_0! k_1! \ldots k_m!} a_0^{k_0} a_1^{k_1} \cdots a_m^{k_m}.$$

By factoring out a_0^{s-1} from column s, we can regain the denominator $(-1)^m a_0^{2m+1}$ of (32):

$$b_4 = \frac{1}{a_0^{15}} \begin{vmatrix} c_0 & a_0 & & & \\ c_1 & a_1 & a_0^2 & & \\ c_2 & a_2 & 2a_0a_1 & a_0^3 & \\ c_3 & a_3 & 2a_0a_2 + a_1^2 & 3a_0^2a_1 & a_0^4 \\ c_4 & a_4 & 2a_0a_3 + 2a_1a_2 & 3a_0^2a_2 + 3a_0a_1^2 & 4a_0^3a_1 \end{vmatrix} =$$

$$= \frac{1}{a_0^9} \begin{vmatrix} c_0 & a_0 & & & \\ c_1 & a_1 & a_0 & & \\ c_2 & a_2 & 2a_1 & a_0 & \\ c_3 & a_3 & 2a_2 + a_0^{-1}a_1^2 & 3a_1 & a_0 \\ c_4 & a_4 & 2a_3 + 2a_0^{-1}a_1a_2 & 3a_2 + 3a_0^{-1}a_1^2 & 4a_1 \end{vmatrix}. \quad (33')$$

In the last formula, $k_0 \leq 0$ and $\sum k_j = 1$ at each entry. Sometimes, presentation (33') can be simplified. For instance, for the inverse problem, we have

$$b_0 = \frac{1}{a_0}, \quad b_1 = \frac{-1}{a_0^3}a_1, \quad b_2 = \frac{1}{a_0^5}\begin{vmatrix} a_1 & a_0 \\ a_2 & 2a_1 \end{vmatrix}, \quad b_3 = \frac{-1}{a_0^7}\begin{vmatrix} a_1 & a_0 & \\ a_2 & 5a_1 & a_0 \\ a_3 & 4a_2 & a_1 \end{vmatrix},$$

$$b_4 = \frac{1}{a_0^9}\begin{vmatrix} a_1 & a_0 & & \\ a_2 & 2a_1 & a_0 & \\ a_3 & 8a_2 & 7a_1 & a_0 \\ a_4 & 5a_3 & 3a_2 & a_1 \end{vmatrix}, \qquad b_5 = \frac{-1}{a_0^{11}}\begin{vmatrix} a_1 & a_0 & & & \\ a_2 & 2a_1 & a_0 & & \\ a_3 & 7a_2 & 7a_1 & a_0 & \\ a_4 & 13a_3 & 14a_2 & 3a_1 & a_0 \\ a_5 & 6a_4 & 6a_3 & a_2 & a_1 \end{vmatrix}. \quad (33'')$$

a much more pleasing form than (33'). Although this kind of presentation with positive integer coefficients is not even unique for b_4 and b_5, it does not seem to exist for b_6 at all.

Our guide in the rest of the section is Sir Thomas Muir's four volume book [Mui11]. This is a list of results of articles written about determinants in the last century. Proofs are not usually given, the reader is referred to the original articles.

It was Spottiswoode in 1853, who found the following formula for the k^{th} derivative of the quotient of two functions:

$$\left(\frac{g}{f}\right)^{(k)} = (-1)^k \frac{\det M_{k+1}}{f^{k+1}},$$

$$\text{where } M_{k+1} = \begin{bmatrix} g & f & & & \\ g' & f' & \ddots & & \\ \vdots & & \ddots & \ddots & \\ g^{(k-1)} & & \ddots & \ddots & f \\ g^{(k)} & \binom{k}{k} f^{(k)}(x) & \cdots & \binom{k}{2} f''(x) & \binom{k}{1} f'(x) \end{bmatrix}. \quad (34)$$

Using Taylor's formula and (34), we can obtain

$$\left(\frac{g(x)}{f(x)}\right)^{(k)} = \frac{1}{f(0)} \sum_{k=0}^{\infty} \frac{1}{(-f(0))^k} \frac{\det M_{k+1}(0)}{k!} x^k, \text{ provided } f(0) \neq 0. \quad (35)$$

Twenty five years later Glaisher realized that if $f(x) = \sum_{k=0}^{\infty} a_k x^k$ $(a_0 \neq 0)$ and $g(x) = \sum_{k=0}^{\infty} b_k x^k$, then $\frac{\det M_{k+1}(0)}{k!}$ in (35) can be simplified (we gave a proof in Lemma 4 of Section 3.):

$$\frac{\sum_{k=0}^{\infty} b_k x^k}{\sum_{k=0}^{\infty} a_k x^k} = \frac{1}{a_0} \left(1 + \sum_{k=1}^{\infty} \frac{\det M_k}{(-a_0)^k} x^k\right),$$

$$\text{where } M_k = \begin{bmatrix} b_0 a_1 - b_1 a_0 & a_0 & & & \\ b_0 a_2 - b_2 a_0 & a_1 & \ddots & & \\ \vdots & & \ddots & \ddots & \\ b_0 a_{k-1} - b_{k-1} a_0 & & & \ddots & a_0 \\ b_0 a_k - b_k a_0 & a_{k-1} & \cdots & \cdots & a_1 \end{bmatrix}. \quad (36)$$

Observe that M_k in (36) reduces to (26) if we want to calculate the reciprocal of the power series $f(x)$.

Once we can compute $\frac{f'(x)}{f(x)}$ using (36), we can integrate it to obtain $\log f(x)$ and then $e^{f(x)}$:

$$\log\left(1+\sum_{k=1}^{\infty} a_k x^k\right) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{\det M_k}{k} x^k, \tag{37}$$

and

$$e^{\sum_{k=0}^{\infty} a_k x^k} = e^{a_0}\left(1+\sum_{k=1}^{\infty}\frac{\det M_k}{k!}x^k.\right). \tag{38}$$

In (37) and (38) the matrices M_k are of type (27) (with $a_0 := 1$) and (28) (with $a_0 := -1, f_j := j$) respectively.

Let us conclude with H. W. Segar's beautiful result (1892) from [Muil1, vol. IV., page 236] about rational powers of a power series:

$$\left(\sum_{k=0}^{\infty} a_k x^k\right)^{\frac{n}{m}} = a_0^{\frac{n}{m}}\left(1+\sum_{k=1}^{\infty}\frac{\det M_k}{k! m^k a_0^k} x^k\right),$$

where $a_0 \neq 0$ and $M_k =$

$$\begin{bmatrix} na_1 & -ma_0 & & & \\ 2na_2 & (n-m)a_1 & -2ma_0 & & \\ \vdots & & \ddots & \ddots & \\ (k-1)na_{k-1} & & & \ddots & -(k-1)ma_0 \\ kna_k & ([k-1]n-m)a_{k-1} & \dots & (2n-[k-2]m)a_2 & (n-[k-1]m)a_1 \end{bmatrix}. \tag{39}$$

Notice that the only non-integer exponent (branch point) is at the leading $a_0^{\frac{n}{m}}$ factor.

REFERENCES

[And76] G. E. Andrews, "The Theory of Partitions," Addison-Wesley, Reading, MA, 1976.

[And88] G. E. Andrews and F. G. Garvan, *Dyson's crank of a partition*, Bull. Am. Math. Soc. **18** (1988), 167–171.

[Aga87] A. K. Agarwal, G. E. Andrews, D. M. Bressoud, *The Bailey lattice*, J. Indian Math Soc. **51** (1987), 57–73.

[Ait56] A. C. Aitken, "Determinants and Matrices," Oliver and Boyd, Edinburgh, Scotland, 1956.

[Atk54] A. O. L. Atkin, H. P. F. Swinnerton-Dyer, *Some properties of functions*, Proc. London Math. Soc. **(3) 4** (1954), 84–106.

[Atk67] A. O. L. Atkin, *Proof of a conjecture of Ramanujan*, Glasgow Math J. **8** (1967), 14–32.

[Bar93] M. Barnsley and L. Hurd, "Fractal Image Compression," AK Peters Ltd., Wellesley, MA, 1993.

[Bre83] D. Bressoud, *A matrix inverse*, Proc. Amer. Math. Soc. **88** (1983), 446–448.

[Com74] Louis Comtet, "Advanced Combinatorics," Reidel, Dordrecht, Holland, 1974.

[Dav66] F. N. David, M. G. Kendall and D. E. Barton, "Symmetric Function and Allied Tables," University Press, Cambridge, England, 1966.

[Egg89] P. C. Eggan, *Linear congruences for power products of Dedekind η-functions*, J. London Math. Soc. **(2) 39** (1989), 231–245.

[Ego77] G. P. Egorichev, "Integral Representation and the Computation of Combinatorial Sums," AMS Translations of Math. Monographs, 1984 (1977).

[Fin88] N. Fine, "Basic Hypergeometric Series and Applications," Amer. Math. Soc., Providence, RI, 1988.

[Gar90] F. Garvan, D. Kim and D. Stanton, *Cranks and t-cores*, Invent. Math. **101** (1990), 1–17.

[Gar93] F. Garvan, *Congruences for partitions that are p-cores*, Proc. London Math. Soc. **(3) 66** (1993), 449–478.

[Gas90] G. Gasper and M. Rahman, "Basic Hypergeometric Series," Cambridge U. Press, 1990.

[Gel71] A. O. Gelfond, "Calculus of Finite Differences," Hindustan Publ. Corp. (transl.), 1971.

[Gou61] H. W. Gould, *The q-Stirling Numbers of First and Second Kinds*, Duke Math. J. **28** (1961).

[Ham70] S. J. Hammarling, "Latent Roots and Latent Vectors," U. of Toronto Press, 1970.

[Har92] J. Harris, "Algebraic Geometry: a First Course," Springer Verlag, 1992.

[Hir92] Philip S. Hirschorn and Louise Arakeljan Raphael, *Coalgebraic foundations of the method of divided differences*, Advances in Math. **91** (1992), 75–135.

[Jam81] G. James and A. Kerber, "The Representation Theory of the Symmetric group," Addison-Wesley, Reading MA, 1981.

[Kly82] A. A. Klyachko, *Modular forms and representations of symmetric groups*, (English transl.) Zapiski Instituta im. V. A. Steklova (Leningrad) **116** (1982).

[Knu92] Donald E. Knuth, *Two Notes on Notation*, Amer. Math. Monthly (1992).

[Knu93] Donald E. Knuth, *Johann Faulhaber and Sums of Powers*, Math. Comp. **61** (1993).

[Kri86] V. Krishnamurty, "Combinatorial Theory and Applications," Ellis Horwood, Chichester, England, 1986.

[Lak90] S. Lakshmivarahan and S. K. Dhall, "Analysis and Design of Parallel Algorithms," McGraw-Hill, 1990.

[Lil93] G. M. Lilly and S.C. Milne, *The C_ℓ Bailey transform and Bailey lemma*, J. Constr. Approx. **9** (1993), 473–500.

[Mac79] I. G. MacDonald, "Symmetric Functions and Hall Polynomials," Clarendon Press, Oxford, England, 1979.

[Mel73] Z. A. Melzak, "Companion to Concrete Mathematics," John Wiley and Sons, New York, NY, 1973.

[Mui11] Thomas Muir, "The Theory of Determinants in the Historical Order of Development, Vol. I-IV.," Macmillan and Co., London, England, 1911-23.

[Mui28] Thomas Muir, "A Treatise on the Theory of Determinants," Dover Publ., 1960, revised by W. Metzler in 1928, originally published in 1883.

[Niv91] I. Niven, "An Introduction to the Theory of Numbers," Wiley, New York, NY, 1991.

[Ono93] Ken Ono, "Congruences on the Fourier coefficients of modular forms on $\Gamma_0(N)$ with number-theoretic applications," PhD. dissertation, U. of California, Los Angeles, 1993.

[Pol37] G. Pólya and R. C. Read, "Combinatorial Enumeration of Groups, Graphs and Chemical Compounds," Springer, 1937, new edition 1987.

[Rit91] G. Ritter, *Heterogenous matrix products*, Proceedings of SPIE **1568** (1991).

[Rom78] S. Roman and G.-C. Rota, *The Umbral Calculus*, Adv. Math. **27** (1978).

[Wat38] G. N. Watson, *Ramanujans Vermutung über Zerfällungsanzahlen*, J. Reine Angew. Math. **179** (1938), 97–128.

BIOGRAPHICAL SKETCH

Zoltán Réti was born on December 5, 1956, in Budapest, Hungary. He received a mathematics degree from Lóránd Eötvös University, Budapest in 1981 and a Master's degree in mathematics from University of South Carolina in 1989.

He worked in the Computer and Automation Institute of the Hungarian Academy of Sciences from 1981 to 1986.

His research interests include number theory, combinatorics, linear algebra, and approximation theory.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Krishnaswami Alladi

Krishnaswami Alladi, Chairman
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Frank Garvan

Frank Garvan
Associate Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Li-Chien Shen
Associate Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arun Kumar Varma

Arun Varma
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Malay Ghosh

Malay Ghosh
Professor of Statistics

This dissertation is submitted to the Graduate Faculty of the Department of Mathematics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1994

Dean, Graduate School